



**HAL**  
open science

# Epidémiologie moléculaire et évolution de l'entérovirus A71 et interactions génétiques avec les autres entérovirus de l'espèce A responsables de la maladie pied-main-bouche.

Chervin Hassel

► **To cite this version:**

Chervin Hassel. Epidémiologie moléculaire et évolution de l'entérovirus A71 et interactions génétiques avec les autres entérovirus de l'espèce A responsables de la maladie pied-main-bouche.. Médecine humaine et pathologie. Université d'Auvergne - Clermont-Ferrand I, 2015. Français. NNT : 2015CLF1MM03. tel-01366694

**HAL Id: tel-01366694**

**<https://theses.hal.science/tel-01366694>**

Submitted on 15 Sep 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ BLAISE PASCAL

UNIVERSITÉ D'Auvergne

Année 2015

N° d'ordre

*ÉCOLE DOCTORALE  
DES SCIENCES DE LA VIE ET DE LA SANTÉ, AGRONOMIE,  
ENVIRONNEMENT*

*Thèse*

Présentée à l'Université d'Auvergne, Clermont-Ferrand  
pour l'obtention du grade de (Décret du 5 juillet 1984)

**DOCTEUR D'UNIVERSITÉ**

**Spécialité : BIOLOGIE CELLULAIRE et MOLECULAIRE**

Le 21 avril 2015 par

**Chervin HASSEL**

---

**ÉPIDÉMIOLOGIE MOLECULAIRE ET ÉVOLUTION  
DE L'ENTEROVIRUS A71 ET INTERACTIONS  
GÉNÉTIQUES AVEC LES AUTRES ENTEROVIRUS  
DE L'ESPECE A RESPONSABLES DE LA  
MALADIE PIED-MAIN-BOUCHE**

---

Directeur de Thèse :  
**Monsieur le Docteur Jean-Luc BAILLY**

Devant la commission d'examen composée de

Examineurs : Monsieur le Docteur Jean-Luc BAILLY  
Monsieur le Docteur Téléphore SIME-NGANDO

Rapporteurs : Monsieur le Professeur Bruno POZZETTO  
Monsieur le Docteur Francis DELPEYROUX

**Laboratoire de Virologie - équipe EPIE - EA4843**

Centre National de Référence entérovirus/parechovirus laboratoire associé



# REMERCIEMENTS

Mes remerciements les plus chaleureux aux membres du jury qui m'ont fait l'honneur d'accepter de juger mon travail de thèse. Toute ma gratitude et mes remerciements à Mr Télésphore Sime-Ngando le président du jury et à mes rapporteurs, Monsieur Francis Delpeyroux et Monsieur Bruno Pozzetto d'avoir consacré du temps à l'analyse et l'évaluation de ce travail scientifique. Un grand Merci pour vos précieux conseils et vos encouragements.

Je remercie tout particulièrement mon directeur de thèse, Monsieur Jean-luc Bailly, responsable de l'EA4843 – laboratoire Epidémiologie et pathogénie des infections à entérovirus à l'Université d'Auvergne. Il a su me faire découvrir et me faire aimer l'univers de la virologie médicale depuis le stage de Master 2. Cette thèse témoigne de son excellent travail de supervision qu'il a réalisé avec beaucoup de patience, de gentillesse, de bonne humeur, de disponibilité et surtout d'humanité. Ses précieux conseils m'ont permis de me guider et d'apprendre considérablement et d'acquérir surtout la rigueur et le savoir-faire scientifique qui me seront nécessaires dans ma future carrière. Je ne t'exprimerai jamais assez toute ma gratitude pour tout ce qu'il m'a appris et pour l'expérience acquise à tes côtés.

Ma profonde gratitude à Madame Hélène Peigue-Lafeuille pour m'avoir accueilli au sein du laboratoire de virologie, pour ses précieux conseils et pour son soutien tout le long de ma thèse.

Un grand merci à Madame Audrey Mirand avec qui j'ai eu le plaisir de travailler sur le projet du protocole pied-main-bouche. Tes compétences, tes précieux conseils et ton aide m'ont permis d'avancer efficacement mon projet de thèse.

Mes remerciements les plus chaleureux à Madame Cécile Henquell, Madame Christine Archimbaud et Madame Martine Chambon pour leur gentillesse, leur soutien et leurs conseils.

Je tiens à remercier Madame Françoise Charbonnet pour ses précieux conseils et son aide dans l'initiation de la culture de la lignée HMEC-1.

Un grand merci à Isabelle Simon pour m'avoir initié aux cultures cellulaires, son aide très précieuse dans la réalisation de ce projet, ses conseils, son humour et sa gentillesse.

Merci à Gwendoline Jugie pour sa précieuse aide en biologie moléculaire, son aide fut fondamentale à l'avancement de ce projet, sa grande patience avec les PCR, ses conseils, son implication et sa gentillesse.

Mes sincères remerciements à Nathalie Rodde pour son aide, ses conseils, sa gentillesse, sa patience, son expérience et son implication.

Mes chaleureux remerciements à Romain Volle avec qui j'ai eu le plaisir de travailler pendant trois ans. Ton humour, tes encouragements et ton aide m'ont été très précieux. Cette complicité au quotidien au labo et lors des congrès m'a beaucoup apporté.

Mes sincères et chaleureux remerciements à Maxime Bisseux avec qui j'ai eu le plaisir de travailler au quotidien. Ton amitié compte beaucoup pour moi et je te souhaite tout le meilleur pour la suite sur le plan personnel et professionnel.

Merci à Amélie Brebion pour son aide et ses conseils et à qui je souhaite toute la meilleure réussite durant sa thèse.

Mes remerciements à Xavier Despont, stagiaire de Master 1 MGEB avec qui j'ai travaillé sur la partie recombinaison ainsi que Valentin Simonin, stagiaire de DUT ABB avec qui j'ai travaillé sur la partie des cellules HMEC-1. Je vous souhaite de la réussite dans votre avenir professionnel.

Une pensée naturelle à tous mes anciens camarades stagiaires du laboratoire de Virologie, entre autres, Jérémie Azevedo, Sébastien Lhomme, Céline Ribière et Mathias Raymond.

Je tiens à remercier l'ensemble des personnels du laboratoire de Virologie du Centre de Biologie (secrétaires, techniciennes, stagiaires) pour leur aide et leur accueil.

Un grand merci à à Hervé, Rémi, Magalie ainsi que l'ensemble de l'équipe de DSI de la Faculté de Médecine-Pharmacie. Merci pour votre « patience », votre aide, votre humour et vos compétences.

Je tiens à remercier particulièrement Monsieur Antoine Mahul du CRRI pour ses compétences et son aide logistique pour la réalisation des analyses bayésiennes.

Un remerciement chaleureux à Christelle Balvignac et Lorraine Gameiro du CICS pour leur aide et soutien.

Un grand Merci à l'ensemble du personnel du laboratoire de Madame Christiane Forestier et notamment à Damien et à Nicolas pour leur aide.

Durant ma thèse, j'ai eu la chance d'obtenir une activité d'enseignement au sein de l'équipe pédagogique du laboratoire de Biophysique Neurosensorielle de la Faculté de Pharmacie, dirigé par Monsieur Michel Doly. Un grand merci à Monsieur David Cia qui m'a formé dans le domaine particulier de l'enseignement de l'informatique et du C2i. J'ai eu un grand plaisir d'apprendre à tes côtés et tu m'as transmis la passion du métier d'enseignant chercheur.

Un grand remerciement chaleureux à toute l'équipe Isabelle Ranchon-Cole, Françoise Gachon, Sylviane Rousselin, Nathalie Jacquemont, Christine Cergy et Khaoula. Ce travail d'équipe m'a permis de poursuivre ma carrière à vos côtés par le biais d'un poste d'ATER. Un grand merci à Isabelle pour ton soutien et tes encouragements dans cette voie.

Merci à « ERTICa team » pour leur aide, leur soutien et leurs encouragements durant ces trois années de thèse : Mahchid Bamdad, Emmanuelle Moutenou, Pierre Daumar, Corinne Aubel, Carole Zanin, Robin Dufour et Clémence Dubois.

Merci à tous mes proches de m'avoir soutenu, Thibaud, Agnes, Louis, Stephanie, Jérémy, votre amitié est précieuse pour moi.

Un immense merci à mes parents, à ma famille tout particulièrement à ma grand-mère et une pensée pour une personne partie dans un monde meilleur.



# TABLE DES MATIERES

<b>Lexique</b> .....	<b>10</b>
<b>Abréviations</b> .....	<b>13</b>
<b>Introduction générale</b> .....	<b>15</b>
<b>Caractéristiques générales des entérovirus</b> .....	<b>18</b>
I. Classification et taxonomie des entérovirus .....	19
II. Structure de la capsidite des entérovirus .....	20
III. Génome viral .....	21
1) Région 5' non codante .....	21
2) Cadre ouvert de lecture .....	22
3) Région 3' non codante .....	22
IV. Cycle viral intracellulaire .....	22
1) Les récepteurs d'entrée dans la cellule hôte et la décapsidation.....	22
2) La traduction du génome viral .....	23
3) Réplication du génome viral .....	24
4) Assemblage des virions.....	25
5) Libération des virions .....	25
6) Effets de la multiplication virale sur la cellule hôte.....	25
A. Modulation de l'expression des gènes cellulaires pendant l'infection.....	26
B. Réorganisation des structures intracellulaires .....	26
V. Les molécules servant de récepteurs cellulaires.....	27
1) Les récepteurs de la super famille des immunoglobulines .....	27
2) Les récepteurs de la famille des intégrines .....	28
3) Le récepteur DAF (Decay Accelerating Factor, CD55) .....	28
VI. Réplication virale et diversité génétique des entérovirus .....	28
1) Mutations ponctuelles .....	28
2) Recombinaison génétique .....	29
A. Recombinaison homologue dite « choix de copie ».....	30
B. Recombinaison non homologue « non réplivative » .....	30
VII. Principales pathologies et diagnostic des entérovirus .....	31
1) Atteintes neuroméningées .....	31
2) Autres pathologies .....	31
3) Diagnostic .....	32
4) Identification moléculaire (génotypage) .....	32
<b>L'entérovirus type A71 et les sérotypes apparentés de l'espèce A</b> .....	<b>34</b>
I. Manifestations cliniques.....	35



1)	Maladie pied-main-bouche .....	35
2)	Herpangine .....	36
3)	Paralysie flasque de type poliomyélite .....	36
4)	Encéphalite du tronc cérébral.....	36
II.	Pouvoir pathogène.....	37
III.	Récepteurs cellulaires de l'EV-A71 et des types apparentés .....	38
1)	PSGL-1 (P-selectin glycoprotein ligand-1) .....	39
2)	SCARB2 (human scavenger receptor class B, member 2) .....	39
3)	L'annexine II.....	40
4)	L'acide sialique .....	41
5)	La lectine DC-SIGN « DC-specific intercellular adhesion molecule-3 grabbing nonintegrin » .....	41
IV.	Modèles animaux de l'infection à EV-A71 .....	41
1)	Modèle simien.....	41
2)	Modèle « rongeur » .....	42
V.	Epidémiologie des infections à EV-A.....	43
1)	Epidémiologie de l'EV-A71 .....	43
A.	Souche prototype.....	43
B.	Premières vagues épidémiques, années 1970–1990 [Tableau 4] .....	43
C.	Epidémies dans la région Asie Pacifique depuis 1997.....	44
D.	Epidémies et circulation de l'EV-A71 en Europe .....	46
2)	Epidémiologie des infections causées par les autres EV-A .....	46
3)	Surveillance de la maladie pied-main-bouche .....	47
A.	Epidémio-surveillance dans les pays de la région Asie-Pacifique.....	47
B.	Epidémio-surveillance des infections à entérovirus aux USA et en Europe.....	48
C.	Epidémio-surveillance des infections à entérovirus en France .....	48
VI.	Séroprévalence de l'infection à EV-A71 .....	49
VII.	Epidémiologie moléculaire des entérovirus de l'espèce A.....	50
1)	Epidémiologie moléculaire de l'EV-A71 .....	50
2)	Epidémiologie moléculaire des coxsackievirus A .....	52
3)	Recombinaison génétique et évolution des entérovirus A.....	53
VIII.	Variabilité génétique et immunité croisée .....	54
IX.	Prévention des infections et développement d'un vaccin.....	55
	<b>Les méthodes d'analyse phylodynamique : concepts généraux et application à l'épidémiologie virale .....</b>	<b>58</b>
I.	Phylogénie moléculaire .....	59
1)	Méthode des distances .....	60
2)	Maximum de parcimonie .....	61

3)	Méthodes probabilistes - Vraisemblance .....	61
4)	Généralités sur l'inférence Bayésienne et son application en phylogénie .....	62
II.	Datation moléculaire .....	64
III.	Coalescence et démographie .....	65
IV.	Phylogéographie .....	66
V.	Evolution adaptative des protéines.....	67
VI.	Le logiciel BEAST .....	68
1)	Estimation de généalogies les plus probables (taux d'évolution et TMRCA).....	68
2)	Estimation de la démographie.....	69
3)	Estimation de la distribution géographique des lignées.....	69
	<b>Objectifs des travaux de la thèse.....</b>	<b>71</b>
1)	Epidémiologie moléculaire de l'EV-A71 et des principaux EV-A responsables du syndrome pied-main-bouche .....	72
2)	Evolution structurale de la capsid de l'EV-A71 .....	72
3)	Rôle de la recombinaison génétique dans l'évolution de l'espèce EV-A .....	72
4)	Etude du pouvoir pathogène des EV-A avec un modèle cellulaire de barrière endothéliale microvasculaire de derme .....	73
	<b>Matériels et méthodes .....</b>	<b>74</b>
I.	Souches d'entérovirus .....	75
1)	Echantillon constitué pour étudier l'épidémiologie moléculaire des EV-A .....	75
2)	Echantillon constitué pour étudier l'évolution de la capsid de l'EV-A71 .....	75
3)	Echantillon constitué pour étudier la recombinaison chez les EV-A.....	75
4)	Echantillon constitué pour étudier le pouvoir infectieux des EV-A .....	75
5)	Isolement des souches et titrage du pouvoir infectieux .....	75
6)	Désignation des souches .....	76
II.	Extraction de l'ARN viral et synthèse d'ADNc.....	76
III.	Amplification par RT-PCR de différents fragments génomiques .....	76
1)	Séquence génique 1D <sup>VP1</sup> .....	76
2)	Région génomique P1 .....	77
3)	Séquence génique partielle 3CD .....	77
4)	Fragment subgénomique compris entre 1D <sup>VP1</sup> et 3CD .....	77
5)	Génome complet .....	77
IV.	Purification et séquençage des produits d'amplification.....	77
V.	Constitution des jeux de séquences nucléotidiques (Tableau 10).....	78
1)	Analyses préalables des jeux de séquences.....	78
2)	Jeux des séquences 1D <sup>VP1</sup> .....	78
3)	Jeu des séquences P1 constitué pour l'EV-A71 .....	78
4)	Jeux des séquences 3CD .....	79

5)	Jeux des séquences subgénomiques et des génomes complets .....	79
VI.	Analyse de la diversité des EV-A .....	79
VII.	Analyse exploratoire de la recombinaison .....	79
1)	Recherche de souches d'origine recombinante par différences de topologies.....	79
2)	Analyse de la similarité nucléotidique des subgénomomes / génomes complets .....	80
3)	Analyse des signatures moléculaires de recombinaison .....	80
VIII.	Analyse du signal temporel dans les jeux de séquences .....	80
IX.	Analyse phylogénétique par une méthode bayésienne .....	80
X.	Reconstruction de l'histoire démographique des populations virales .....	81
XI.	Estimation de l'origine géographique des nœuds ancestraux .....	81
XII.	Polymorphisme des protéines de capsid .....	82
1)	Calcul de l'entropie de Shannon .....	82
2)	Reconstruction des états ancestraux.....	82
XIII.	Estimation des pressions de sélection dans les protéines de capsid.....	83
XIV.	Numéros d'accèsion des séquences nucléotidiques .....	83
XV.	Etude de l'interaction des entérovirus avec un modèle de barrière endothéliale de derme 84	
1)	Lignée HMEC-1 (Human Microvascular Endothelial Cell).....	84
2)	Sensibilité de la lignée HMEC-1 à l'infection par les entérovirus de l'espèce A..	84
3)	Modèle <i>in vitro</i> de barrière endothéliale microvasculaire de derme.....	85
<b>Travaux personnels .....</b>		<b>86</b>
I.	Epidémiologie moléculaire comparée des sérotypes d'EV-A.....	87
1)	Polymorphisme des séquences moléculaires et diversité intratypique .....	87
2)	Epidémiologie moléculaire de l'EV-A71 .....	88
A.	Evaluation de l'horloge moléculaire dans les jeux de données de l'EV-A71 ...	88
B.	Profils phylogénétiques des populations d'EV-A71 reliées à l'épidémiologie..	88
C.	Histoire démographique des populations d'EV-A71 en Europe et en Asie .....	91
D.	Phylogéographie des populations européennes d'EV-A71 .....	92
3)	Epidémiologie moléculaire des coxsackievirus A .....	94
II.	Evolution moléculaire de la capsid de l'EV-A71 .....	96
1)	Phylogénie moléculaire de l'EV-A71 établie avec les séquences de la région P1	96
2)	Variation d'entropie dans les séquences protéiques P1 de l'EV-A71 .....	98
3)	Estimation des substitutions non synonymes dans la généalogie de l'EV-A71 ...	98
4)	La sélection à l'échelle moléculaire parmi les populations d'EV-A71 .....	99
III.	Recombinaison génétique et diversité des sérotypes de l'espèce EV-A .....	100
1)	Recherche des souches d'EV-A d'origine recombinante – Jeux de séquences ...	100
A.	Première phase de l'étude : analyse phylogénétique exploratoire .....	100

B.	Séquençage génomique et constitution des jeux de données .....	103
2)	Analyse des signatures d'événements de recombinaison génétique.....	103
A.	Analyse des séquences subgénomiques .....	103
B.	Analyse des séquences de l'ORF du CV-A16.....	106
C.	Analyse des séquences de l'ORF de l'EV-A71.....	106
D.	Analyse croisée des séquences de l'ORF du CV-A16 et de l'EV-A71 .....	107
3)	Datation des évènements de recombinaison par analyses bayésiennes .....	107
IV.	Etude du pouvoir pathogène des EV-A avec un modèle cellulaire de barrière endothéliale microvasculaire de derme .....	108
1)	Mise au point des conditions de culture de la lignée HMEC-1 .....	108
2)	Sensibilité de la lignée HMEC-1 à l'infection par différents EV-A.....	109
3)	Développement d'un modèle de barrière endothéliale de derme.....	111
<b>DISCUSSION</b>	.....	<b>113</b>
I.	L'épidémiologie virale analysée à partir de données moléculaires.....	114
II.	Etudes antérieures de l'histoire évolutive de l'EV-A71.....	114
III.	Apport de la dynamique géographique à l'épidémiologie de l'EV-A71.....	115
IV.	Dynamique des populations virales et épidémiologie des autres EV-A.....	117
V.	Apport de la phylodynamique à l'épidémiologie des maladies virales.....	118
VI.	Evolution adaptative de la capsid de l'EV-A71 au cours de la circulation des souches .....	119
VII.	Recombinaison génétique et évolution des EV-A.....	120
VIII.	Développement d'un modèle d'étude du pouvoir pathogène des EV-A .....	122
<b>Références Bibliographiques</b>	.....	<b>124</b>
<b>Annexes</b>	.....	<b>139</b>

# LEXIQUE

**Akaike Information Criterion (AIC)** : Un critère de sélection basé sur le maximum de vraisemblance, avec lequel il est possible de comparer une hypothèse (dans le cadre de l'étude, un modèle d'évolution moléculaire) qu'elle soit niché ou non. L'hypothèse qui présente la valeur AIC la plus petite est considérée comme la plus probable. Le test statistique AIC est un compromis entre l'adéquation du modèle aux données (mesurée par la valeur de vraisemblance  $\ln$ ) et le nombre de paramètres du modèle car les modèles les plus riches en paramètres auront nécessairement une meilleure adéquation aux données.

**Coalescence** : Processus décrivant la rencontre de deux lignées en un même ancêtre lorsque l'on parcourt en sens inverse du temps la généalogie d'un échantillon de gènes dans une population.

**Datation moléculaire** : La datation moléculaire consiste à coupler un arbre phylogénétique décrivant les relations de parentés entre des taxons, reconstruits à partir de séquences d'ADN ou de protéines, avec des points de calibration (ou d'étalonnage) temporels afin d'estimer les âges absolus de divergence entre des organismes. Un arbre phylogénétique est inféré à partir de séquences d'un gène et les longueurs de branches sont calculées/déduites des changements observés dans les séquences. La quantité d'évolution allant de la racine à chacune des feuilles est variable. Cependant, de telles variations sont probablement dues à l'aspect stochastique de l'évolution moléculaire. Lorsque ces différences de vitesse d'évolution ne sont pas statistiquement significatives, l'arbre peut être converti en une phylogénie ultra métrique qui correspond à une parfaite horloge moléculaire globale. Si on dispose d'une information pour étalonner cet arbre (l'âge de l'ancêtre commun de deux taxons), on peut en déduire le taux de substitutions nucléotidiques par unité de temps et l'âge des autres nœuds de l'arbre phylogénétique (horloge moléculaire globale).

**Démographie** : Etude quantitative des populations et leurs dynamiques dans le temps et dans l'espace, à partir de données telles que les effectifs, la structure en taille ou en âge, la natalité, la fécondité, la variabilité ou la migration. Dans notre étude, la démographie des populations virales est analysée, modélisée à partir des effectifs, plus précisément de la taille efficace.

**Généalogie** : Relations phylogénétiques estimées à partir de séquences échantillonnées dans le temps dont la racine et les nœuds internes sont datés.

**Horloge moléculaire** : Taux d'évolution apparemment constant à long terme, ce qui n'exclut pas la possibilité de fluctuations autour d'une valeur moyenne à court terme. Les horloges moléculaires se comptent plus de manière stochastique que selon un rythme régulier

- Importance des points de calibration, du choix du gène
- Cas des gènes qui n'évoluent pas à taux constants (il existe des tests de détection des écarts par rapport à l'hypothèse d'horloge moléculaire). Par exemple, en utilisant un cadre probabiliste, il est possible de savoir si l'arbre avec HM est significativement moins vraisemblable que l'arbre sans HM. L'application d'un taux d'évolution erroné (HM mal étalonnée) conduit à d'importantes erreurs dans l'estimation des âges de divergence, en particulier les anciens.

**Incongruence** : L'incongruence désigne des différences significatives entre les histoires évolutives inférées pour un même groupe d'espèces. Une collection d'arbres phylogénétiques est dite incongruente si les différentes histoires évolutives qu'ils modélisent présentent de fortes dissemblances. Une collection de marqueurs phylogénétiques est dite incongruente si les arbres phylogénétiques inférés à partir de chacun d'entre eux forment une collection d'arbres.

**Inférence** : L'inférence consiste à induire les caractéristiques inconnues d'une population à partir d'un échantillon issu de cette population. Les caractéristiques de l'échantillon, une fois connues, reflètent avec une certaine marge d'erreur possible celles de la population.

**Jeux de données** : Alignement de séquences

**Lignée** : Une lignée correspond à l'ensemble d'entités (dans notre étude, des souches virales) possédant une racine commune. Une lignée phylogénétique correspond à l'ensemble des taxons dérivant d'un taxon donné.

**Modèle (statistique)** : Objet mathématique, basé sur des hypothèses permettant d'analyser les données, c'est-à-dire de tirer des conclusions à partir de données.

**Monophylétique** : Se dit d'un groupe naturel, c'est-à-dire reflétant l'histoire évolutive des organismes. Un ensemble de taxons est dit monophylétique lorsque celui-ci regroupe tous les descendants d'un même ancêtre commun direct, et uniquement entre eux.

**Population** : Ensemble d'individus d'une même espèce coexistant dans un même habitat. Par extension, dans cette étude, une population virale désigne un ensemble de souches ayant des caractéristiques épidémiologiques et génétiques communes.

**Recombinaison** : Processus par lequel l'information génétique contenue dans deux chromosomes homologues est réorganisée pour aboutir à la formation de nouveaux chromosomes, constituant un mélange de chromosomes parentaux. Par extension, dans cette étude, nous utilisons le terme de recombinaison pour désigner tout processus qui associe l'information génétique issue de génomes viraux différents.

La recombinaison est dite intertypique lorsque les souches virales dont les génomes sont recombinaison, appartiennent à des sérotypes différents.

La recombinaison est dite intratypique lorsque les souches virales dont les génomes sont recombinaison, appartiennent aux mêmes sérotypes.

**Sérotipe** : Manière de nommer les subdivisions taxonomiques (de classement) de micro-organismes sur la base des caractéristiques de leur antigène ou protéines.

**Taxon** : Entité conceptuelle censée regrouper tous les organismes vivants possédant en commun certains caractères taxonomiques ou diagnostiques bien définis. La classification taxonomique propose une hiérarchie codifiée en sept rangs principaux : Règne, embranchement, classe, ordre, famille, genre, espèce. Par extension, dans cette étude, nous incluons les rangs hiérarchiques suivants : sérotypes, génogroupes et sous génogroupes, couramment utilisés dans le domaine d'étude des agents infectieux.

**Topologie** : D'un point de vue qualitatif, en cladistique, la topologie décrit l'agencement relatif des nœuds et des branches d'un arbre.

# ABBREVIATIONS

3Dpol : Polymérase 3D  
ADN : Acide désoxyribonucléique  
AIC : critère d'information d'Akaike  
ARN : Acide ribonucléique  
ASR : Ancestral Sequence Reconstruction  
BEAST : Bayesian Evolutionary Analysis Sampling Trees  
BLAST : Basic local alignment search tool  
BSP : Bayesian skyline plot  
BSSVS : Bayesian stochastic search variable selection  
CDC : Centers for Disease Control and Prevention  
CHU : Centre Hospitalier Universitaire  
CNR : Centre National de Référence  
CV : Coxsackievirus  
CV-A16 : Coxsackievirus A16  
ECP : effet cytopathogène  
EV-A71 : Entérovirus A71  
EV-A : Entérovirus de l'espèce A  
FMDV : Foot and mouth disease virus  
GTR : General time reversible  
IRES : Internal ribosome entry site (Site d'initiation interne de la traduction)  
JDD : Jeux de données  
LCR : Liquide céphalo rachidien (liquide cérébro-spinal)  
LRT : Likelihood ratio test  
MCMC : Markov Chain Monte-Carlo  
ML : Maximum likelihood  
MRC-5 : Medical Research Council 5 (cellules de type fibroblastique)  
NC : Non Codant  
NCBI : National Center for Biotechnology Information  
NLS : Nuclear localization site  
ORF : Open reading frame (cadre de lecture ouvert)



Pb : Paire de bases  
PCR : Polymerase chain reaction  
pH : Potentiel hydrogène  
PMB : Pied-main-bouche  
PSGL-1 : P-selectin glycoprotein ligand-1  
PV : Poliovirus  
RNC : région non codante (voir aussi UTR)  
RD : Rhabdomyosarcome (lignée de cellules de type épithélial)  
RE : Réticulum endoplasmique  
RF : Formes répliquatives  
RNC : Région non codante  
RT-PCR : Reverse Transcriptase Polymerase Chain Reaction  
SCARB2 : Scavenger receptor class B, member 2  
SNC : Système Nerveux Central  
SV-40 : Simian virus 40  
TMRCA : Time to the most recent common ancestor  
UTR : Untranslated region  
UV : Ultra-violet

# **INTRODUCTION GENERALE**

La découverte des entérovirus est liée à une maladie épidémique : la poliomyélite. Une stèle de l’Égypte antique (1403-1365 av. JC) semble indiquer que cette maladie était présente depuis très longtemps. C’est au début du 20<sup>ème</sup> siècle que la poliomyélite devient un problème de santé publique important en raison de la survenue de nombreuses épidémies en Europe et aux USA. Les premiers travaux pour rechercher l’origine de la maladie mettent en évidence un agent « ultrafiltrable » en 1904. Puis en 1949, le virus est cultivé pour la première fois *in vitro* en utilisant des tissus embryonnaires humains. Cette avancée a ouvert de nombreuses perspectives, notamment celle de produire de grandes quantités de virus avec une plus grande facilité pour la production de vaccins développés par Salk et Sabin. Parallèlement aux progrès réalisés en matière de physiopathologie et de prévention vaccinale contre la poliomyélite, d’autres virus présentant des propriétés physico-chimiques proches de celles des poliovirus ont été mis en évidence par culture, jusqu’à ce que l’on découvre que le poliovirus faisait partie d’un groupe plus vaste désigné par le terme d’entérovirus, reconnaissant ainsi leur propriété de virus entérique. Ce groupe sera reconnu comme un genre de la famille virale des *Picornaviridae* (pico : petit et rna : acide ribonucléique).

Les picornavirus sont des virus de petite taille dépourvus d’enveloppe et contenant un génome d’ARN monocaténaire de polarité positive. Cette famille virale contient de nombreux agents pathogènes humains et animaux. Selon leurs propriétés biologiques et génétiques, ils ont été classés en huit genres taxinomiques : Aphovirus, Cardiovirus, Entérovirus, Erbovirus, Hepatovirus, Kobuvirus, Parechovirus et Teschovirus (**Fauquet *et al*, 2005 ; Pallansch & Roos, 2007**). Quatre d’entre eux (Entérovirus, Hepatovirus, Parechovirus et Kobuvirus) comprennent des virus pathogènes pour l’homme.

Comme plusieurs autres virus à ARN, certains picornavirus (FMDV, agent de la fièvre aphteuse ; poliovirus, agent de la poliomyélite) font l’objet d’une surveillance particulière. Ils présentent des caractéristiques communes : multiplication intracellulaire intense, taux élevé de mutations permettant une adaptation rapide aux contraintes de l’environnement liées à l’hôte. Ces facteurs conditionnent la transmission interindividuelle des souches, la persistance des populations\* virales au cours du temps et dans l’espace, et l’épidémiologie des maladies associées aux infections. Cette dissémination des virus peut être rapide entre des zones géographiques distantes. Par exemple, celle du coronavirus responsable du syndrome respiratoire aigu sévère (SRAS) durant les années 2000 est liée aux moyens de transports modernes qui contribuent à des déplacements internationaux en très peu de temps (moins de temps que la période d’incubation de certains virus hautement pathogènes). Une épidémie locale peut ainsi menacer d’autres régions du monde. A cela s’ajoute le grand potentiel de

mutations et d'adaptation rapide des virus à ARN. Le maintien d'un processus ininterrompu de transmission virale est lié à la capacité adaptative des souches, de sorte que le devenir d'un mutant dépend de son pouvoir pathogène mais aussi des réactions de défense de l'hôte. Concernant les virus entériques, des facteurs non sélectifs de résistance liés à leur présence dans l'environnement, sont aussi à prendre en considération.

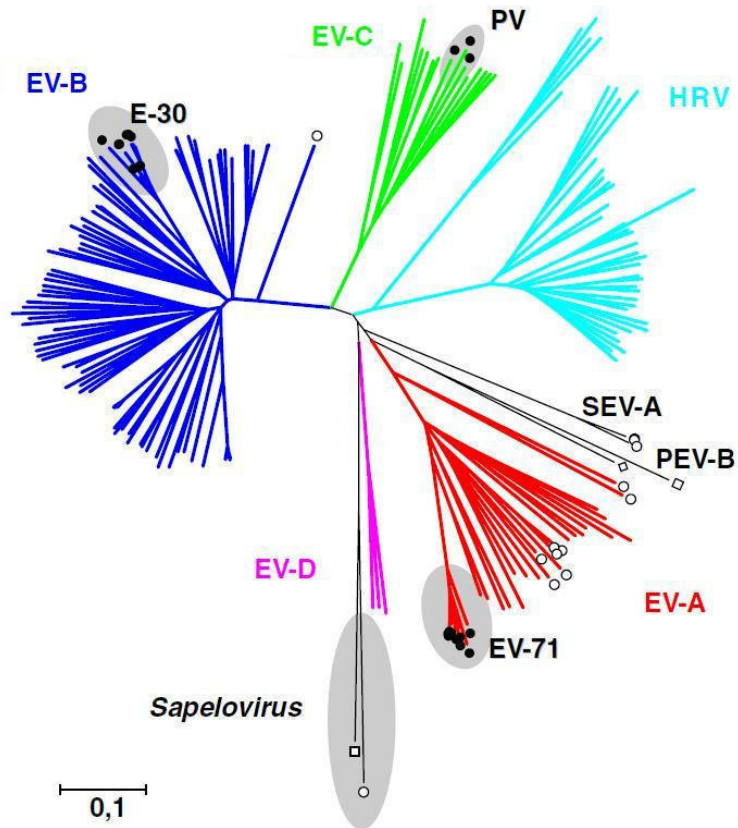
Dans ce contexte d'évolution virale rapide (en temps réel par rapport à l'observateur), les méthodes d'épidémiologie moléculaire sont des outils particulièrement intéressants pour analyser l'histoire évolutive des populations virales, déterminer leurs structures et leur modes d'interaction avec l'hôte (pouvoir pathogène, variation antigénique...). L'épidémiologie moléculaire repose sur les principes de l'analyse phylogénétique qui permet d'estimer des relations de parenté et d'établir un classement hiérarchique des souches (taxonomie). La description des relations génétiques entre des taxons apporte des informations statiques à propos des souches identifiées au cours de l'investigation clinique, alors que le pouvoir pathogène et la transmission d'un agent pathogène est un phénomène dynamique. Pour intégrer cet aspect, les méthodes d'analyse phylodynamique ont été développées. Elles apportent des éléments clés pour comprendre l'histoire évolutive (origine dans le temps, vitesse évolutive, adaptation à l'hôte) et les facteurs qui conditionnent les dynamiques migratoires des agents pathogènes (émergence et diffusion spatiales).

### **Présentation du plan général de la revue bibliographique**

La présentation bibliographique comporte trois parties. Dans la première, ont été rassemblés les aspects généraux, les caractéristiques communes à l'ensemble des entérovirus qui permettent d'éclairer aussi les aspects qui n'ont pas fait l'objet d'une étude précise pour les entérovirus de l'espèce A. La deuxième partie présente l'état des connaissances sur les entérovirus A et plus particulièrement de l'entérovirus type A71 (EV-A71). Enfin, il nous a paru nécessaire de donner quelques indications sur les cadres conceptuels multiples qui ont permis le développement des méthodes d'analyse phylodynamique. Il s'agit par exemple de la fonction de vraisemblance et d'éléments d'analyse Bayésienne dans le domaine statistique, de la théorie de la coalescence dans le domaine de la phylogénie et la génétique des populations. Ces aspects ont été présentés dans la troisième partie de l'exposé bibliographique.



# **CARACTERISTIQUES GENERALES DES ENTEROVIRUS**



**Figure 1. Relations phylogénétiques entre les entérovirus humains**

La reconstruction phylogénétique a été obtenue par l'analyse comparée de la séquence nucléotidique complète du gène 1D (protéine de capsid VP1). Les distances génétiques ont été évaluées par la méthode maximum composite likelihood et la reconstruction a été réalisée par la méthode neighbor-joining. Les entérovirus de l'espèce A (EV-A) sont en rouge, l'espèce B (EV-B) en bleu, l'espèce C (EV-C) en vert et l'espèce D (EV-D) en violet. Les rhinovirus sont représentés en cyan. Abréviations : poliovirus (PV), entérovirus 71 (EV-71), echovirus 30 (E-30), simian enterovirus (SEV), porcine enterovirus (PEV) et human rhinovirus (HRV.). Le Sapelovirus a été utilisé comme groupe externe (**d'après Bailly *et al.*, 2008.**

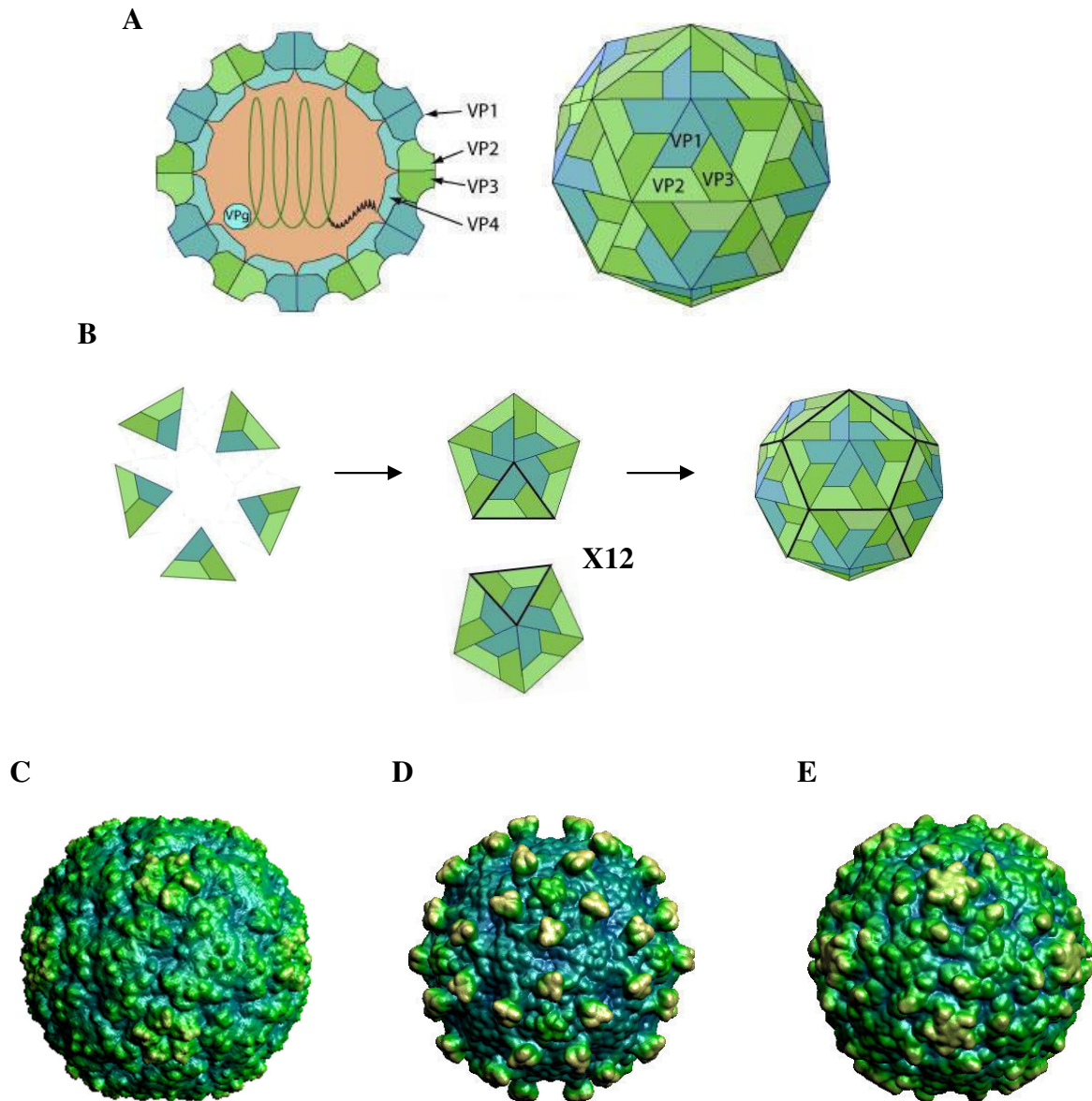
## I. Classification et taxonomie des entérovirus

La première classification des entérovirus a été établie en utilisant les manifestations cliniques et le pouvoir pathogène observé chez des animaux de laboratoire. Ces caractéristiques ont permis de reconnaître les quatre groupes suivants : poliovirus, coxsackievirus A, coxsackievirus B et echovirus (**Committee on the Enteroviruses, 1957**). Les poliovirus sont les agents à l'origine des paralysies flasques humaines, la poliomyélite. Les coxsackievirus ont été isolés pour la première fois à partir de fèces de deux enfants paralysés lors d'une épidémie de poliomyélite. Ils ont été classés en deux sous-groupes, A et B, suivant leur pouvoir pathogène chez le souriceau nouveau-né. Les coxsackievirus A provoquent une paralysie flasque en raison de leur multiplication dans les muscles squelettiques et cardiaques. Les coxsackievirus B présentent un tropisme tissulaire plus large incluant notamment le système nerveux central et ils causent des paralysies spasmodiques. Les echovirus ont été isolés par culture cellulaire alors que leur lien avec la méningite aseptique n'était pas encore établi. Ces virus ont été distingués des autres entérovirus par l'absence de pathologie chez le souriceau. Les souches d'entérovirus ont été différenciées en sérotypes avec les tests de séroneutralisation et classées dans ces quatre groupes (**Lim et Benyesh-Melnick, 1960**). Au début des années 1990, le groupe des entérovirus comprenait trois poliovirus (PV-1 à PV-3), 21 coxsackievirus A (CV-A1 à CV-A14, CV-A16, CV-A17, CV-A19 à A22 et CV-A24), six coxsackievirus B (CV-B1 à B6) et 26 echovirus. Cette classification n'a pas tenu et a subi des adaptations profondes en raison de l'imprécision de la technique de séroneutralisation et de la difficulté de classer des souches toujours plus nombreuses, identifiées par les techniques moléculaires. Les nouveaux sérotypes\* ont reçu la dénomination EV et un numéro (ordre de leur mise en évidence). Plus fondamentalement, une classification basée sur des caractéristiques de parenté génétique a été adoptée.

Depuis les années 2000, le génotypage établi avec les séquences génomiques codant les protéines de capsidie remplace la séroneutralisation pour identifier les souches. Le génotypage est effectué le plus fréquemment avec le gène 1D qui code la protéine de capsidie VP1, ce qui permet une identification précise des souches présentes dans les prélèvements cliniques. Par ailleurs, les caractéristiques phylogénétiques ont permis le classement des sérotypes en espèces taxinomiques : les entérovirus humains ont été inclus dans quatre espèces différentes désignées HEV-A à D [Figure 1].

Les dernières évolutions de la classification ont été rendues nécessaires à cause de deux incohérences majeures. Premièrement, la classification des rhinovirus dans un genre





**Figure 2. Structure et formation de la capside des entérovirus**

Structure d'un entérovirus (A). Les entérovirus sont constitués d'une capsid virale de symétrie icosaédrique qui contient le génome viral. La capsid est constituée de 12 protomères de quatre protéines de structure (VP1 – VP4). Les protéines VP1, VP2, VP3 sont situées à la surface de la capsid alors que VP4 est en position interne. (Source : **Swiss Institut of Bioinformatique 2008**).

Formation de la capsid virale des entérovirus (B). L'assemblage débute par la formation d'un protomère (VP1 – VP4). 5 protomères peuvent se combiner pour former un pentamère et 12 pentamères forment une capsid virale complète. (Source : **Swiss Institut of Bioinformatique 2008**).

Représentation de la surface de la capsid virale de l'Entérovirus A71 (C), Coxsackievirus B3 (D) et du Poliovirus-1 (E).

(Source : VIPERdb <http://viperdb.scripps.edu/>)

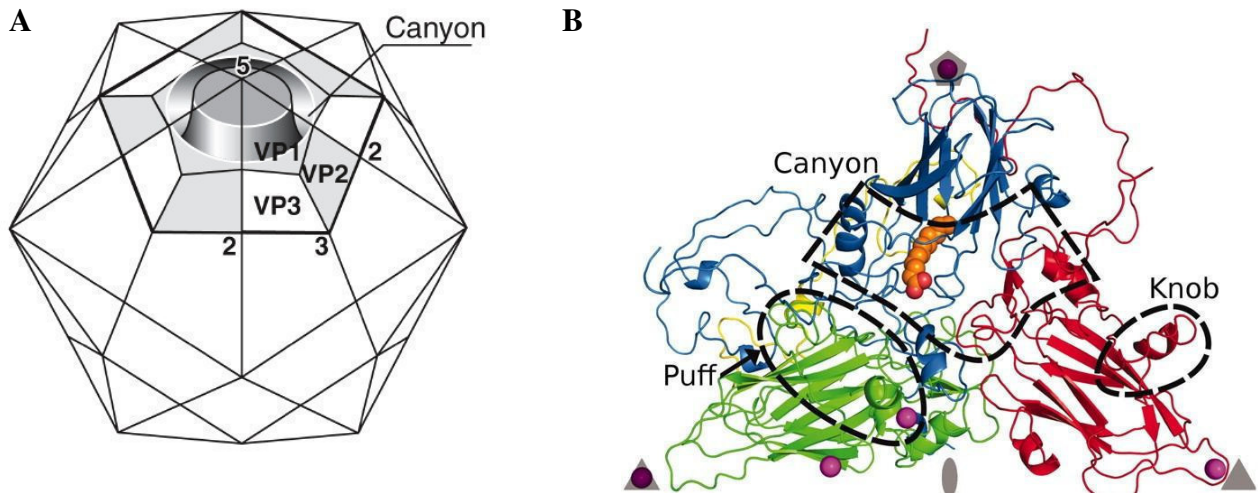
séparé de celui des entérovirus était en contradiction avec les critères de distance génétique montrant l'existence d'un seul genre. Les rhinovirus ont donc été inclus dans le genre unique Entérovirus tout en conservant leur désignation habituelle. La seconde incohérence était liée à la dénomination des espèces taxinomiques qui faisait référence à une seule espèce hôte, l'homme. Par exemple, la désignation de l'espèce HEV-A impliquait qu'elle contenait uniquement des virus d'origine humaine, ce qui était en contradiction objective avec les résultats des analyses phylogénétiques indiquant que de nombreux entérovirus isolés chez différents singes étaient apparentés à des entérovirus humains. La désignation actuelle des espèces taxinomiques d'entérovirus a donc perdu la référence à une espèce hôte. Actuellement, le genre entérovirus comprend environ trois cents sérotypes répartis en douze espèces taxinomiques ; les quatre espèces dans lesquelles se retrouvent les virus pathogènes pour l'homme sont désignées EV-A à EV-D [Figure 1].

## II. Structure de la capsid des entérovirus

Les virions des entérovirus sont sphériques (diamètre d'environ 30 nm) et chaque particule virale comprend une molécule de génome et une capsid protéique [Figure 2a]. Ces virus ne présentent pas d'enveloppe lipidique, ce qui leur confère une résistance aux solvants des lipides (alcool à 70°, éther et chloroforme) et aux acides faibles (sauf les Rhinovirus). La conformation compacte de la capsid protège le génome viral contre des facteurs environnementaux délétères (température, pH). Cependant, les entérovirus sont thermolabiles à partir de 50°C et sensibles au formol, au glutaraldéhyde, à l'hypochlorite de sodium (Javel) et aux acides forts [Melnick, 1996 ; Chambon *et al*, 2004].

La capsid virale est constituée de l'assemblage de 60 exemplaires des protéines VP1, VP2, VP3 et VP4 [Figure 2a]. Les quatre protéines forment une sous unité ou protomère (Rueckert *et al*, 1969), 5 protomères se combinent pour former un pentamère et 12 pentamères constituent la capsid [Figure 2b]. L'analyse des capsides par cristallographie aux rayons X a révélé la structure tridimensionnelle de plusieurs entérovirus. Malgré la diversité des quatre protéines structurales parmi les entérovirus, l'architecture globale de la capsid virale demeure conservée [Figures 2c, 2d, 2e].

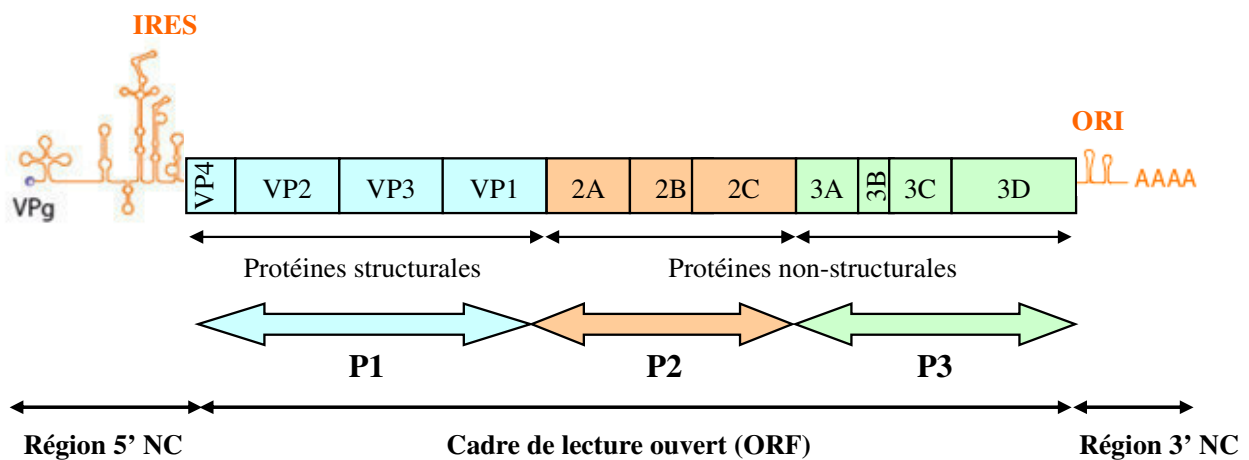
Les protéines VP1, VP2 et VP3 sont en partie exposées à la surface des virions et leur topologie semblable est constituée de 8 feuillets  $\beta$  antiparallèles dont le repliement constitue une forme dite en tonneau [Figure 3a]. Ces trois protéines sont reliées entre elles par des



**Figure 3. Structure de la surface de la capsid des entérovirus**

Topologie d'un protomère d'entérovirus 71 (A). Les protéines de structure du protomère, VP1, VP2, VP3 et VP4, sont représentées respectivement en bleu, vert, rouge et jaune. Les pointillés indiquent les positions identifiables à la surface de la capsid dont le canyon au niveau de la protéine VP1 (d'après Plevka *et al.*, 2012).

Structure de la surface de la capsid des entérovirus (B). La structure tridimensionnelle d'un protomère (VP1-VP4) est représentée. Ce dernier est délimité par un axe de symétrie 5, un axe de symétrie 3 et de 2 axes de symétrie 2. La protéine VP1 présente une dépression qui forme un canyon se prolongeant tout le long d'un pentamère (d'après Blondel *et al.*, 2006).



**Figure 4. Structure du génome des entérovirus**

Le génome des entérovirus est composé d'un seul cadre ouvert de lecture (ORF) d'environ 6,5 kb encadré par deux régions non codantes (5'NC et 3'NC). L'ORF est composé des régions génomiques P1, P2 et P3 qui codent les différentes protéines virales structurales (P1) et non structurales (P2 et P3). Les régions non codantes possèdent des structures secondaires dues aux repliements de la molécule d'ARN, impliquées dans la traduction (IRES en 5') et la réplication du génome viral (ORI en 3'). La région 5'NC peut lier la protéine virale VPg, protégeant ainsi la molécule d'ARN viral.

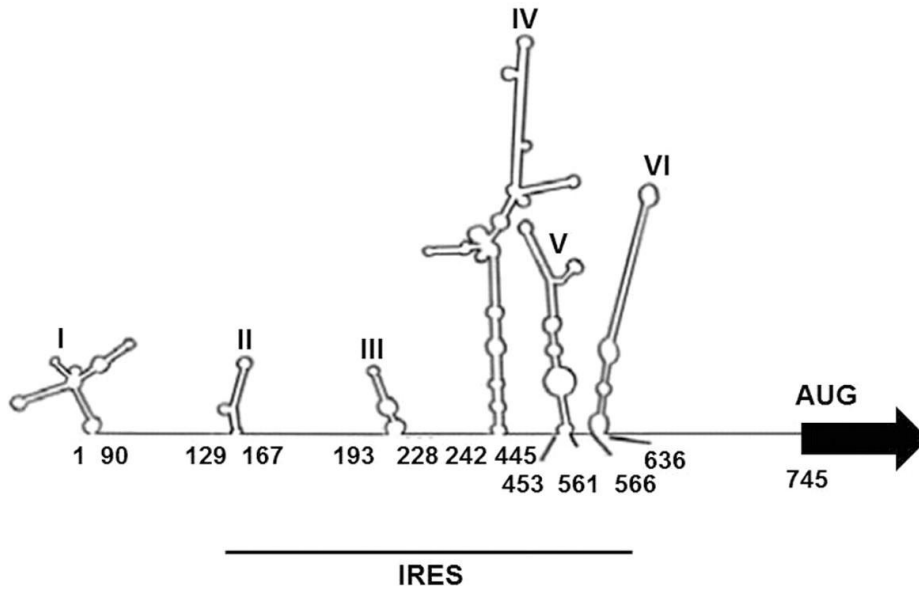
interactions de leur domaine N-terminal et également par la protéine VP4, localisée à l'intérieur de la capsid. Les trois protéines VP1, VP2 et VP3 présentent des zones particulièrement variables et exposées à la surface des virions. Il s'agit de boucles peptidiques qui se projettent à la surface de la capsid et portent les épitopes caractéristiques de chaque sérotype (Minor *et al*, 1986). Selon le sérotype, la taille et la conformation des boucles procurent des variations de relief à la capsid virale. Les cinq protomères d'un même pentamère se rejoignent par la protéine VP1 autour de l'axe d'ordre 5 et des boucles de cette protéine forment un plateau entouré d'une dépression appelée « canyon » [Figure 3b]. Cette fosse abrite les sites de fixations aux récepteurs cellulaires des poliovirus et des rhinovirus (Belnap *et al*, 2000) ; elle est donc reliée à des fonctions importantes du cycle infectieux et au pouvoir pathogène. Des boucles proéminentes sont formées dans les autres protéines majeures : pour la protéine VP2, il s'agit du « puff » et pour la protéine VP3, du « knob » ; ces structures peuvent aussi être impliquées dans la liaison aux récepteurs [Figure 3a].

### **III. Génome viral**

Le génome des entérovirus est constitué d'une molécule d'ARN non segmenté (environ 7500 nucléotides), simple brin et de polarité positive ; il est donc traduit directement lorsqu'il atteint le cytoplasme de la cellule hôte. L'organisation fonctionnelle est conservée chez tous les entérovirus : un seul cadre ouvert de lecture (ORF) d'environ 6,5 kb encadré par les régions non codantes (RNC) en 5' et 3' [Figure 4] (Racaniello, 2007).

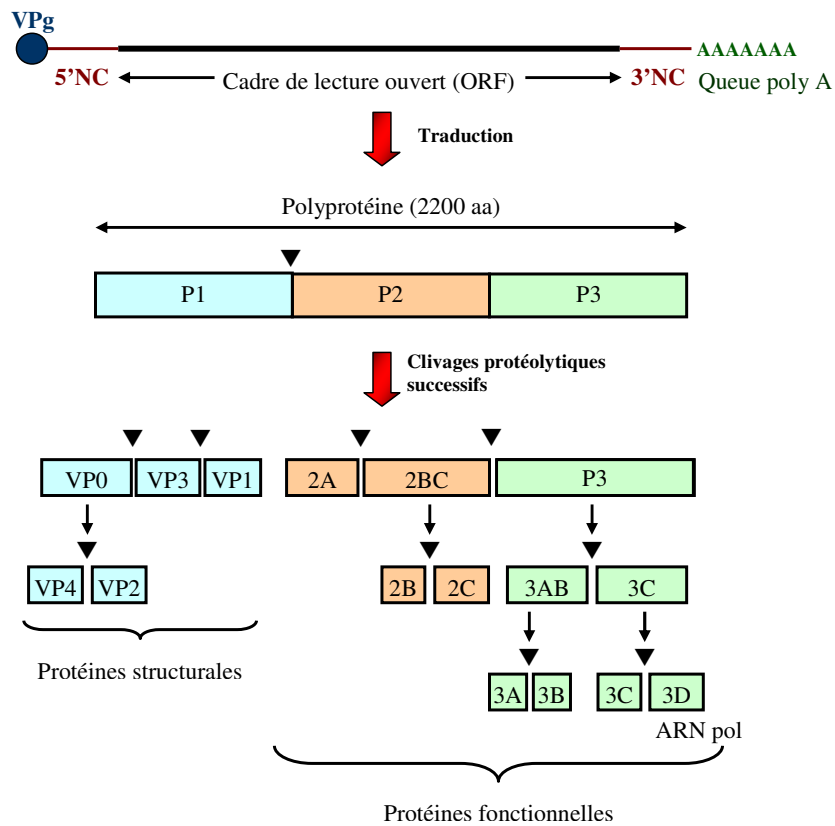
#### **1) Région 5' non codante**

Le repliement de la molécule d'ARN correspondant à la RNC 5' génère six domaines ou structures secondaires (succession de structures tige-boucle) conservées chez tous les entérovirus [Figure 5]. La première de ces structures secondaires contient un élément multifonctionnel Ori L et participe à la réplication et à la traduction de l'ARN viral par le biais d'interactions ARN-protéines (Gamarnik et Andino, 1998). En aval de l'élément Ori L, se trouve un site d'initiation interne de la traduction (IRES). Celui-ci est composé de cinq structures secondaires et peut recruter les protéines cellulaires nécessaires pour initier la traduction de l'ORF du génome viral. Ce mécanisme d'initiation de la traduction est nécessaire à cause de l'absence de coiffe en 5' de l'ARN viral et de la localisation interne du codon initiateur (positions 740 – 750, selon les sérotypes d'EV).



**FIGURE 5. Représentation schématique des structures secondaires de la séquence 5' non codante du génome de l'EV-A71**

Les 6 structures secondaires en tige boucle ont été prédites à partir du logiciel MFold et sont notées de I à VI. Le premier et le dernier nucléotide de chaque structure ainsi que le codon initiateur (AUG) sont indiqués. La région d'initiation interne de la traduction (IRES) est soulignée (d'après Lin JY *et al.*, 2009).



**FIGURE 6. Synthèse et clivage des protéines virales**

Le génome viral est traduit en une longue polyprotéine de 2200 acides aminés qui contient trois régions : P1, P2 et P3. Elle sera immédiatement clivée en onze protéines par les protéases virales, 2A et 3C.

## 2) Cadre ouvert de lecture

Le long cadre ouvert de lecture (ORF) code une polyprotéine de 2200 acides aminés (250 kDa) dont la maturation est assurée par les protéases virales (clivages post traductionnels) et génère des protéines virales fonctionnelles (Summers et Maizel, 1968) [Figure 6]. L'ORF est organisé en trois régions fonctionnelles (P1, P2 et P3) qui correspondent aux trois premiers précurseurs protéiques produits à partir de la polyprotéine. La région P1 est dite « structurale » car elle code les protéines de la capsid virale VP1, VP2, VP3 et VP4. Les régions P2 et P3 dites « non structurales » codent sept autres protéines virales (2A<sup>pro</sup>, 2B, 2C, 3A, 3B (VPg), 3C<sup>pro</sup> et 3D<sup>pol</sup>) impliquées dans la réplication du génome, la maturation de la polyprotéine et des interactions avec des facteurs intracellulaires. L'ORF contient une structure secondaire dans la séquence génique de la protéine 2C. Ce repliement local de l'ARN viral en forme de tête d'épingle est appelé élément CRE.

## 3) Région 3' non codante

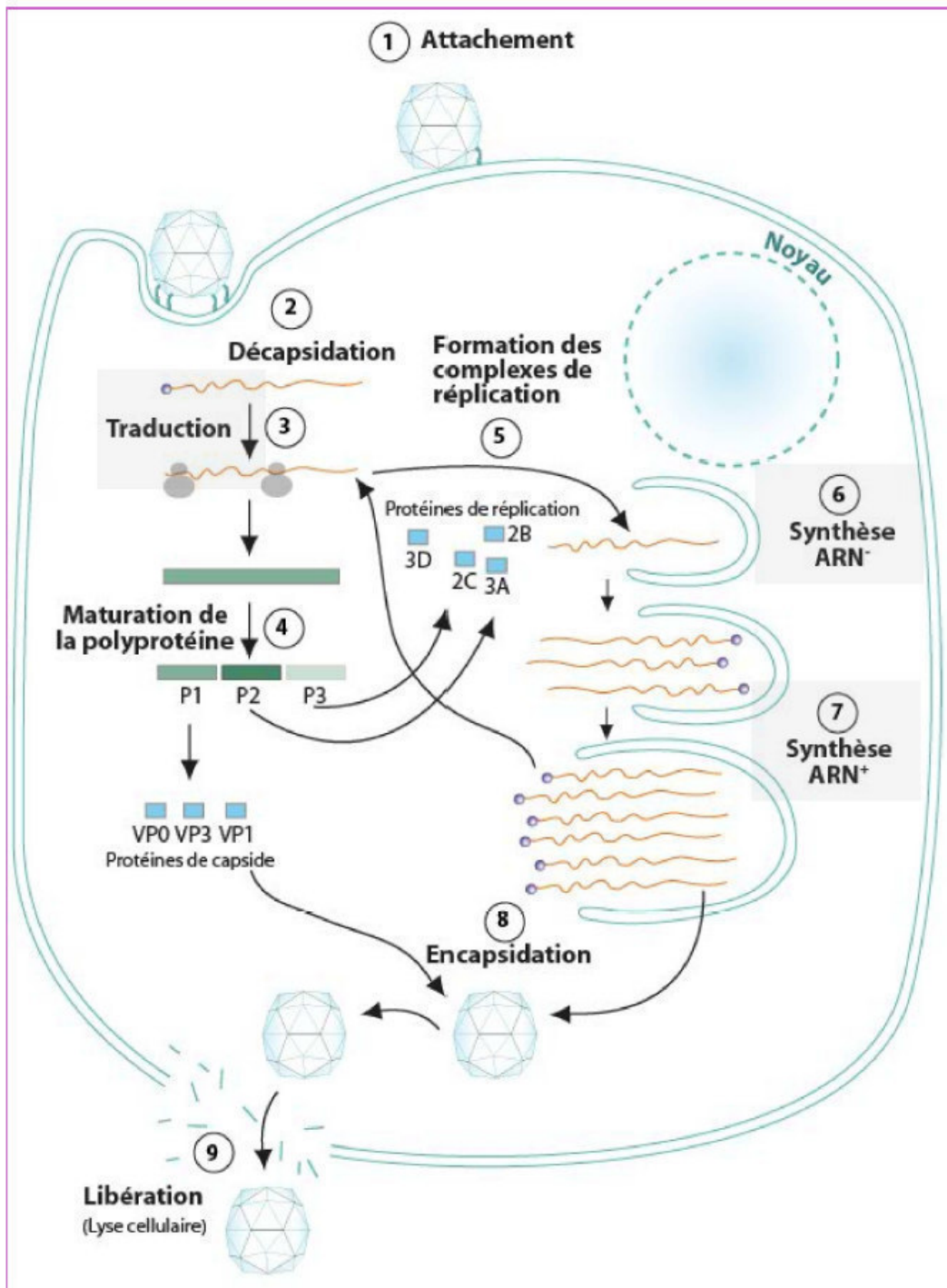
La RNC 3' termine le génome ; elle précède une queue poly adénylée dont la taille est variable (Yogo et Wimmer, 1975). Elle est structurée en domaines tige-boucle, mais contrairement à la RNC 5', les domaines de la R3'NC s'associent en une structure tertiaire de l'ARN dite « pseudo-nœud » qui forme un élément *cis* régulateur Ori R. Cet élément joue un rôle dans l'initiation de la synthèse de l'antigénome (ARN de polarité négative) lors de la réplication virale et son interaction avec l'élément Ori L de la RNC 5' produit une circularisation du génome viral.

# IV. Cycle viral intracellulaire

Le cycle de multiplication des entérovirus se déroule exclusivement dans le cytoplasme de la cellule hôte [Figure 7]. Les aspects principaux sont décrits ci-dessous.

## 1) Les récepteurs d'entrée dans la cellule hôte et la décapsidation

L'entrée du génome dans la cellule hôte s'effectue après l'attachement du virion à un récepteur cellulaire qui interagit avec des sites spécifiques de la capsid, généralement localisés dans le « canyon ». Il existe un grand nombre de récepteurs pour les entérovirus (Cf paragraphe V page 24). L'interaction de la capsid avec un récepteur conduit à des



**FIGURE 7. Cycle viral des Picornavirus**

L'infection est amorcée par l'attachement de la capsidie viral à un récepteur cellulaire (1), entraînant une décapsidation (2) et une translocation de l'ARN viral vers le cytoplasme cellulaire. L'ARN viral est tout d'abord traduit en une polyprotéine (3) dont la maturation génère l'ensemble des protéines virales structurales et non structurales (4). En parallèle, le génome viral sera répliqué à partir d'une première étape de synthèse d'ARN « - » au niveau de complexes de réplication (5,6). Cette ARN sert ensuite de matrice pour la synthèse de nouveaux brins d'ARN « + » (7). Ces ARN « + » serviront à donner de nouvelles protéines virales, qui pourront participer à l'assemblage de nouvelles particules virales (8). Ces provirions seront enfin libérés dans le milieu extracellulaire via une lyse de la cellule hôte (9) pour aller infecter de nouvelles cellules hôtes et réamorcer le cycle viral.

changements de conformation de la particule virale. Le processus d'internalisation par endocytose par récepteur interposé a été analysé en détail pour le poliovirus 1 et le coxsackievirus B3.

Les conséquences immédiates de la fixation du poliovirus à son récepteur CD155 ou hPVR sont des changements de conformation de la particule virale et la libération du génome viral dans le cytoplasme de la cellule hôte (Brandenburg *et al*, 2007). Les modifications conformationnelles résultent de l'externalisation de la protéine VP4 et de la partie N-terminale de la protéine VP1 qui s'insèrent dans la membrane plasmique cellulaire. Un canal hydrophile se forme et assure la translocation de l'ARN viral vers le cytoplasme soit dès le stade de l'invagination de la membrane cellulaire, soit au stade des vésicules d'endocytose.

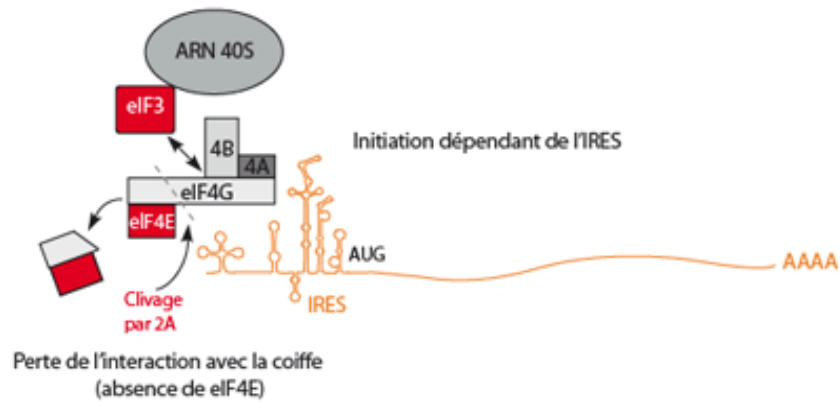
Un autre modèle a été proposé pour décrire l'initiation du cycle infectieux du poliovirus dans les cellules endothéliales microvasculaires cérébrales : l'infection mobilise un mécanisme d'endocytose dépendant des clathrines. La liaison virus/récepteur induit une cascade de signalisation intracellulaire, des réarrangements de l'actine du cytosquelette et l'activation d'une tyrosine kinase essentielle à l'endocytose (Coyne *et al*, 2007).

Le modèle proposé pour expliquer l'entrée du coxsackievirus B3 dans les cellules épithéliales polarisées met en jeu le corécepteur DAF (decay accelerating factor). La fixation du virus au corécepteur active les tyrosines kinases Abl et Fyn et provoque des réarrangements du cytosquelette. Ceci favorise l'interaction du complexe DAF/virion avec le récepteur CAR et induit les changements conformationnels de la capsid virale nécessaires à l'entrée du virus et à la libération du génome viral (Coyne et Bergelson, 2006).

## **2) La traduction du génome viral**

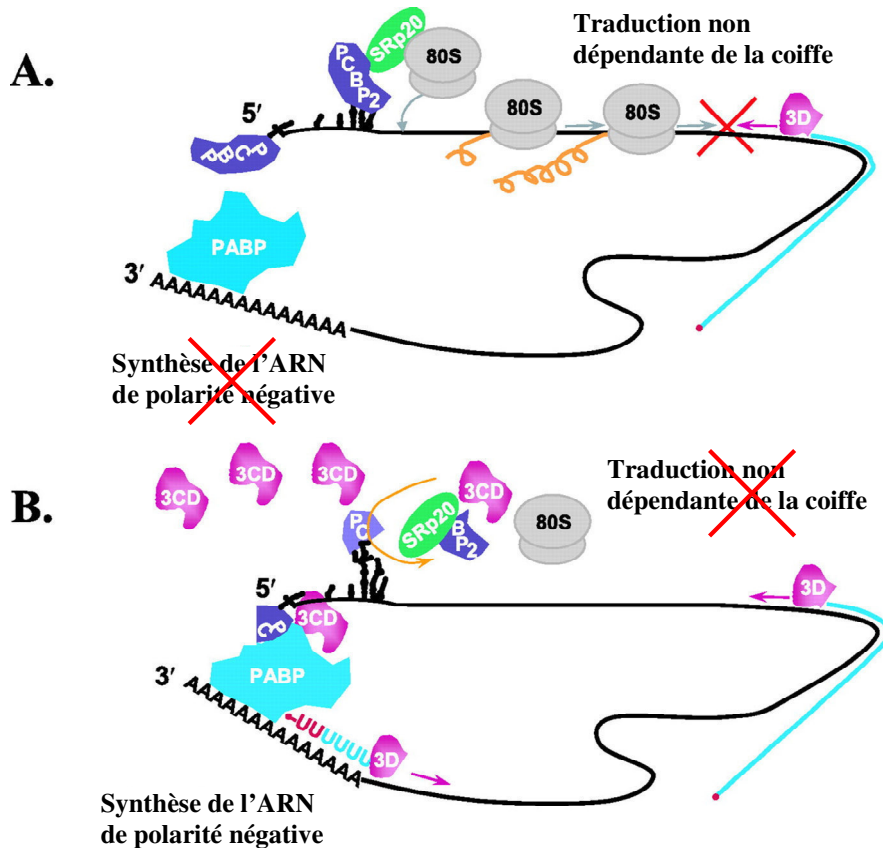
Dès son arrivée dans le cytoplasme de la cellule hôte, le génome viral est libéré par clivage de la protéine VPg via une activité « VPg unlinkase » réalisée par la protéine cellulaire DNA repair enzyme, 5'-tyrosyl-DNA phosphodiesterase-2 (TDP2 ; Virgen-Slane *et al*, 2012). Une fois libérée, il peut être traduit. A la différence de la majorité des ARNm cellulaires, le génome des entérovirus ne contient pas de coiffe pour initier la traduction. A la place, l'élément IRES dans la R5'NC permet l'initiation interne de la traduction (Pelletier et Sonenberg, 1988). Ce mécanisme coiffe-indépendant repose sur la capacité de l'IRES à recruter les ribosomes notamment les sous unités 40S via la protéine PCBP (Gamarnik et Andino, 2000) [Figures 8 et 9]. Le complexe formé parcourt le génome viral jusqu'au codon d'initiation (Racaniello, 2007). La traduction de l'ORF est réalisée d'un seul tenant. Les





**FIGURE 8. Mécanisme moléculaire de la traduction du génome viral des entérovirus**

L'initiation de la traduction est réalisée de manière coiffe indépendante *via* l'IRES situé sur dans la région 5'NC du génome viral. La protéase 2A du virus polio clive le facteur eIF4G, libérant le complexe d'initiation de la traduction (ribosome + eIF4G/4B/4A). Ceci favorise la reconnaissance de l'IRES par ce complexe et initie la traduction du génome viral. La perte du facteur eIF4G limite la traduction des ARNm cellulaires.



**FIGURE 9. Modèle de régulation entre traduction et réplication du génome viral des entérovirus**

Au sein de la cellule hôte, à un instant donné, le génome viral peut être soit traduit de manière coiffe indépendante (A) ou être répliqué par la synthèse d'un brin complémentaire de polarité négative (B). La présence de la protéine 3CD va conditionner la survenue de l'un ou l'autre de ces deux mécanismes.

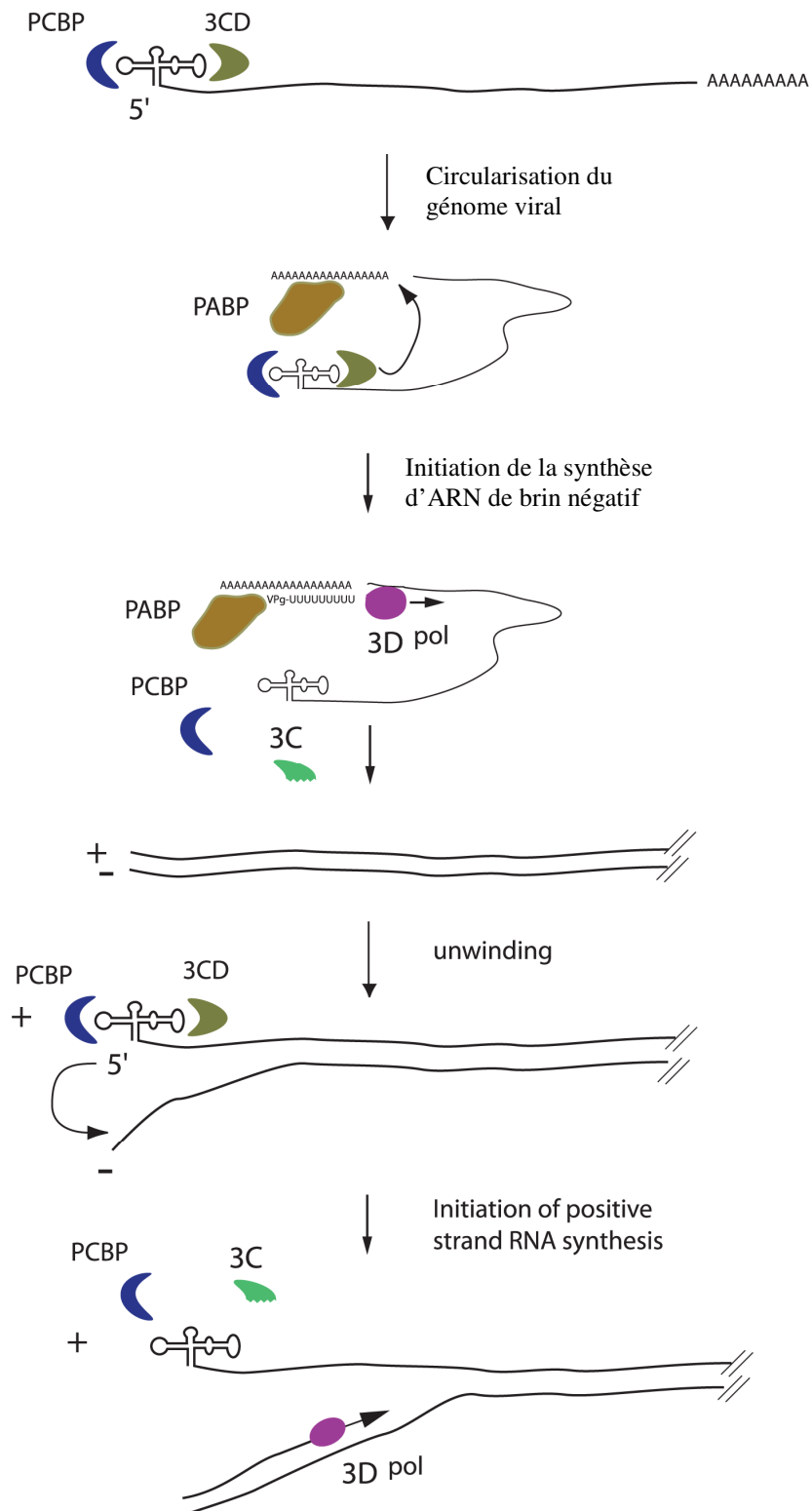
- (A) En l'absence de 3CD, la traduction du génome viral pourra se réaliser *via* le recrutement des ribosomes favorisé par la liaison entre le complexe PCBP2/SRp20/ribosome 80S et la séquence IRES du génome viral.
- (B) En présence de 3CD, cette dernière va cliver la protéine PCBP2, limitant la fixation des ribosomes sur le génome viral. Par ailleurs, la protéine 3CD va interagir avec PABP et PCBP situées respectivement au niveau des régions 5' et 3' non codantes. Ceci entraîne une circularisation du génome viral et favorise la synthèse de l'ARN de polarité négative.

protéines virales sont produites par des clivages de la polyprotéine en cours de synthèse. Les premiers clivages sont autocatalytiques et se produisent dès que les protéases virales sont formées. La protéase 2A<sup>pro</sup> réalise le clivage qui libère le précurseur P1. La seconde protéase 3C<sup>pro</sup> sépare les précurseurs P2 et P3. Chaque précurseur est à nouveau clivé par la protéase 3C<sup>pro</sup> ou son précurseur 3CD<sup>pro</sup>, ce qui génère les protéines virales fonctionnelles (Racaniello, 2007). La protéase 2A<sup>pro</sup> assure le clivage de plusieurs facteurs cellulaires dont le facteur d'initiation de la traduction eIF4G, ce qui inhibe la traduction des ARNm cellulaires dépendant de la coiffe, d'où il résulte une potentialisation de la traduction des ARN viraux (Gradi *et al*, 1998).

### 3) Réplication du génome viral

La réplication du génome des entérovirus repose sur la production d'un ARN de polarité négative qui sert de matrice à la synthèse des génomes viraux de la nouvelle population virale. La transcription des brins complémentaires nécessite la formation d'un complexe de réplication composé de la polymérase virale 3D<sup>pol</sup>, la matrice d'ARN viral et des structures vésiculaires/tubulaires dérivées des membranes du réticulum endoplasmique et du Golgi (Schlegel *et al*, 1986 ; Gamarnik et Andino, 1998). L'action des protéines virales 2BC, 2B, 2C et 3A remodèle les micro-organites cellulaires (Suhy *et al*, 2000). Le réseau de vésicules intriquées qui se constitue, sert de support physique aux complexes de réplication ; il pourrait également protéger les ARN viraux des nucléases (Fogg *et al*, 2003).

La réplication du génome viral ne peut être réalisée en même temps que sa traduction. La régulation entre les deux processus est réalisée par la variation de la concentration de l'intermédiaire 3CD. Cette protéine assure le clivage de PCBP2, un facteur cellulaire nécessaire à la traduction du génome viral, ce qui favorise la réplication du génome viral [Figure 9]. La synthèse du brin de polarité négative met en jeu des interactions protéines-protéines entre le complexe d'initiation (PCBP/3CD) lié au premier domaine de la R5'NC et la protéine PABP (Poly-A Binding Protein) liée à la région 3'NC [Figures 9 et 10]. Ceci entraîne une circularisation du génome viral et permet une poly-uridylation de la protéine VPg par la polymérase virale 3D<sup>pol</sup>. Le complexe d'initiation peut alors amorcer la synthèse du brin de polarité négative. Les hétéroduplexes (ARN double brin contenant un ARN positif et un ARN négatif) constituent les formes répliquatives (RF) ; ils sont dissociés par la fonction hélicase de la protéine virale 2C, ce qui permet de libérer le brin de polarité négative. Le complexe VPg-



**Figure 10. Modèle du mécanisme de répllication génomique des EVs**

Fixation de l'intermédiaire viral 3CDpro et de l'effecteur cellulaire PCBP au niveau de la structure secondaire (5'NC) en feuille de trèfle « Cloverleaf » (I). Formation d'un complexe ternaire *via* la fixation de l'effecteur PABP sur la queue poly A (3'NC), qui permet la circularisation du génome viral et la formation d'un complexe de préinitiation (II). La protéine Vpg après poly uridylation (Vpg-pUpU) va s'hybrider sur la queue poly A et servir d'amorce pour la synthèse de l'antigénome (-) par la polymérase 3Dpol (III). L'ARN bicatenaire est dissocié (IV - V), puis Vpg-pUpU va servir d'amorce sur l'extrémité 3' de l'antigénome (-) pour la synthèse de génomes (+) par 3Dpol (VI) [Vogt et Andino, 2010].

pUpU est recruté et sert aussi d'amorce au niveau de l'extrémité 3' du brin négatif pour initier la synthèse de l'ARN positif par la polymérase virale 3D<sup>pol</sup>. Une molécule matrice peut servir à la synthèse simultanée de six molécules d'ARN de polarité positive.

#### **4) Assemblage des virions**

La morphogénèse des entérovirus procède en plusieurs étapes. Les protéines VP0, VP3 et VP1 produites par le clivage de la région P1 forment un protomère 5S et cinq protomères s'associent en un pentamère 14S (Palmenberg, 1982). Deux hypothèses ont été proposées pour expliquer l'encapsidation du génome viral et la suite du cycle d'assemblage des virions. Le premier modèle de Jacobson et Baltimore (1968) propose que le génome viral s'insère par une ouverture dans une procapside (80S) constituée de douze pentamères produisant un provirion (150S). Le second modèle privilégie un assemblage séquentiel au cours duquel les 12 pentamères s'associent autour du génome viral natif pour former le provirion (Nugent *et al*, 1999). La morphogénèse complète est achevée par le clivage de la protéine VP0 en VP4 et VP2 par un processus encore mal connu. Cette phase génère les particules virales infectieuses (160S) ou virions (Basavappa *et al*, 1994) dans le cytoplasme cellulaire.

#### **5) Libération des virions**

Le mécanisme de libération des virions néo formés hors de la cellule hôte demeure mal connu. Deux modes impliquant ou non la lyse cellulaire ont été décrits. Le premier mécanisme suggère que la nécrose cellulaire est provoquée par la destruction de la membrane plasmique, mais des études récentes suggèrent que l'apoptose ou l'autophagie sont impliqués directement dans la libération des virions (Taylor et Jackson, 2009). Ces deux processus seraient inhibés, évitant l'achèvement de la mort cellulaire avant la fin du cycle viral (VanKuppeveld *et al*, 2005). Les vésicules d'origine autophagique permettraient l'expulsion des virions néo formés dans le milieu extra cellulaire (Jackson *et al*, 2005).

#### **6) Effets de la multiplication virale sur la cellule hôte**

L'environnement cellulaire est profondément modifié pendant le cycle viral ; l'action des protéines virales non structurales dans cette transformation est majeure. Les changements métaboliques et structuraux induisent une morphologie particulière de la cellule infectée,

**Tableau 1 : Fonctions des protéines virales des entérovirus**

<b>Protéine</b>	<b>Fonctions</b>
VP1, VP2, VP3, VP4	Composant de la capside virale
2A	Clivage de protéines virales (protéase) Inhibition de la traduction cellulaire (clivage du facteur de traduction eIF4GI) Blocage du transport nucléo-cytoplasmique par dégradation des complexes des pores nucléaires Induction de l'apoptose par fragmentation de l'ADN nucléaire et clivage de PARP
2B	Altération de la perméabilité membranaire Déséquilibre de l'homéostasie calcique cellulaire Remodelage du Réticulum Endoplasmique et de l'appareil de Golgi.
2C	Rôle dans l'encapsidation (activité NTPase) Rôle dans réarrangement des membranes vésiculaires Blocage de la voie NF- $\kappa$ B
3A	Composant des vésicules membranaires de réplication viral Altération du trafic cellulaire au niveau de l'appareil de Golgi Inhibition de l'apoptose induite par le TNF $\alpha$
3B	Amorce à la réplication du génome viral Associée à l'extrémité 5' du génome viral
3C	Clivage des protéines virales (protéase) Inhibition de la traduction cellulaire (clivage de PABP) Induction de l'apoptose par la voie des caspases
3D	Polymérase à ARN ARN-dépendante Réplication du génome viral Permet l'uridylation de la protéine VPg

l'effet cytopathique (ECP). Les processus cellulaires concernés sont notamment l'expression génique et la réorganisation de structures ou d'organites cytoplasmiques [Tableau 1].

#### **A. Modulation de l'expression des gènes cellulaires pendant l'infection**

La transcription cellulaire est inhibée par l'action de la protéase virale 3Cpro. Cette enzyme migre dans le noyau grâce à un site de localisation nucléaire (NLS). Elle peut cliver de nombreux facteurs d'initiation de la transcription tels la protéine TBP (TATA-binding protein), le facteur de transcription TFIIC et la protéine PABP (Clark *et al*, 1991 ; Clark *et al*, 1993 ; Joachims *et al*, 1999). L'inactivation fonctionnelle de ces facteurs inhibe la transcription effectuée par les ARN polymérase cellulaires de type I, II et III.

La traduction des ARNm cellulaires est inhibée lors de l'infection par un entérovirus notamment par l'action des protéases 3Cpro et 2Apro. Ces deux protéases clivent directement des facteurs impliqués dans l'initiation de la traduction des ARNm coiffés : la protéase 2Apro inactive eIF4GI et eIF4GII (Gradi *et al*, 1998) et la protéase 3Cpro le facteur eIF5B (de Breyne *et al*, 2008). La protéase 2Apro est également reliée à l'altération de composants des pores nucléaires (NPC, Nuclear Pore Complex) dont la conséquence est une limitation du transport nucléo-cytoplasmique des ARNm cellulaires (Park *et al*, 2008).

Les protéines virales 2B et 3A qui disposent de domaines d'ancrage aux membranes lipidiques agissent sur la maturation des protéines. Leur présence dans la cellule est associée à une modification de la perméabilité du réticulum endoplasmique et de l'appareil de Golgi (Doedens *et al*, 1995). La multimérisation de la protéine 2B est reliée à son activité « viroporine » (formation de pores membranaires) qui perturbent l'homéostasie du calcium stocké dans les organites intra cellulaires (Agirre *et al*, 2002) ; ceci provoque un blocage du transport des protéines cellulaires entre le réticulum et de l'appareil de Golgi (DeJong *et al*, 2008).

#### **B. Réorganisation des structures intracellulaires**

La cellule hôte subit de nombreux changements morphologiques lors de l'infection par un entérovirus. Visible en microscopie optique, l'ECP constitue la caractéristique d'un virus lytique : condensation de la chromatine, déformation du noyau, modification de la perméabilité de la membrane cellulaire, changements de l'architecture intracellulaire, fuite des composants intracellulaires... (Bienz *et al*, 1973). La réorganisation complète de la structure intracellulaire comprend la formation de vésicules à multiples membranes, dérivées du RE et de l'appareil de Golgi, qui serviront à la réplication du génome viral (Schlegel *et al*, 1996). L'accumulation des protéines 2B et 3A à la surface ou dans les membranes du RE et

**Tableau 2 : Principaux récepteurs utilisés par les entérovirus pour l'entrée dans la cellule hôte**

<b>Famille</b>	<b>Récepteur</b>	<b>Espèce virale</b>	<b>Sérotypage</b>	<b>Références</b>
<b>Immunoglobuline</b>	PVR (CD55)	EV-C	PV-1, PV-2, PV3	Mendelsohn <i>et al.</i> , 1989
	ICAM-1 (CD54)	EV-C RV	CV-A13, CV-A18, CV-A21 RV	Colonno <i>et al.</i> , 1986 McClelland <i>et al.</i> , 1991 Shafren <i>et al.</i> , 1997a
	CAR	EV-B	CV-B1 – CV-B5	Bergelson <i>et al.</i> , 1997
<b>Intégrine</b>	$\alpha V\beta 3$ (Vitronectine)	EV-B	E-9, CV-A9	Roivainen <i>et al.</i> , 1994 Ylipaasto <i>et al.</i> , 2010
	$\alpha V\beta 6$	EV-B	CV-A9, CV-B1	Agrez <i>et al.</i> , 1997 Heikkilä <i>et al.</i> , 2009
	$\alpha 2\beta 1$ (VLA-2)	EV-B	E-1	Bergelson <i>et al.</i> , 1993
<b>DAF</b>	DAF	EV-B  EV-C EV-D	EV-6, EV-7, EV-11, EV-13, EV-21, EV-29, EV-33 ; CV-B1, CV-B3, CV-B5 ; CV-A21 ; EV-D70	Ward <i>et al.</i> , 1994 Bergelson <i>et al.</i> , 1994 Shieh <i>et al.</i> , 2002 Alexander <i>et al.</i> , 2002
<b>PSGL-1</b>	PSGL-1	EV-A	EV-A71	Nishimura Y <i>et al.</i> , 2009
<b>SCARB2</b>	SCARB2	EV-A	CV-A7, CV-A14, CV-A16, EV-A71	Yamayoshi <i>et al.</i> , 2009
<b>Annexine</b>	Annexine II	EV-A	EV-A71	Yang <i>et al.</i> , 2011
<b>Acide sialique</b>	N-linked Glycan O-linked Glycan	EV-A	EV-A71	Yang <i>et al.</i> , 2009
<b>Lectine</b>	DC-SIGN	EV-A	EV-A71	Lin <i>et al.</i> , 2009

de l'appareil de Golgi est reliée à un processus autophagique avorté, ce qui suggère un détournement de l'autophagie cellulaire par les composants viraux (Suhy *et al*, 2000 ; Taylor et Jackson, 2009). Se forme un réseau dense de vésicules interconnectées qui pourraient également permettre la libération des virions avant la lyse cellulaire (Taylor et Jackson, 2009).

## **V. Les molécules servant de récepteurs cellulaires**

L'évolution des entérovirus sous la forme de multiples sérotypes est intimement liée à l'utilisation de très nombreux récepteurs et corécepteur cellulaires. La diversité des récepteurs explique le vaste tropisme tissulaire des entérovirus. On peut classer les récepteurs des entérovirus selon leurs propriétés physico-chimiques et fonctionnelles [Tableau 2]. Les récepteurs cellulaires des sérotypes de l'espèce EV-A sont présentés dans le chapitre consacré à ces virus.

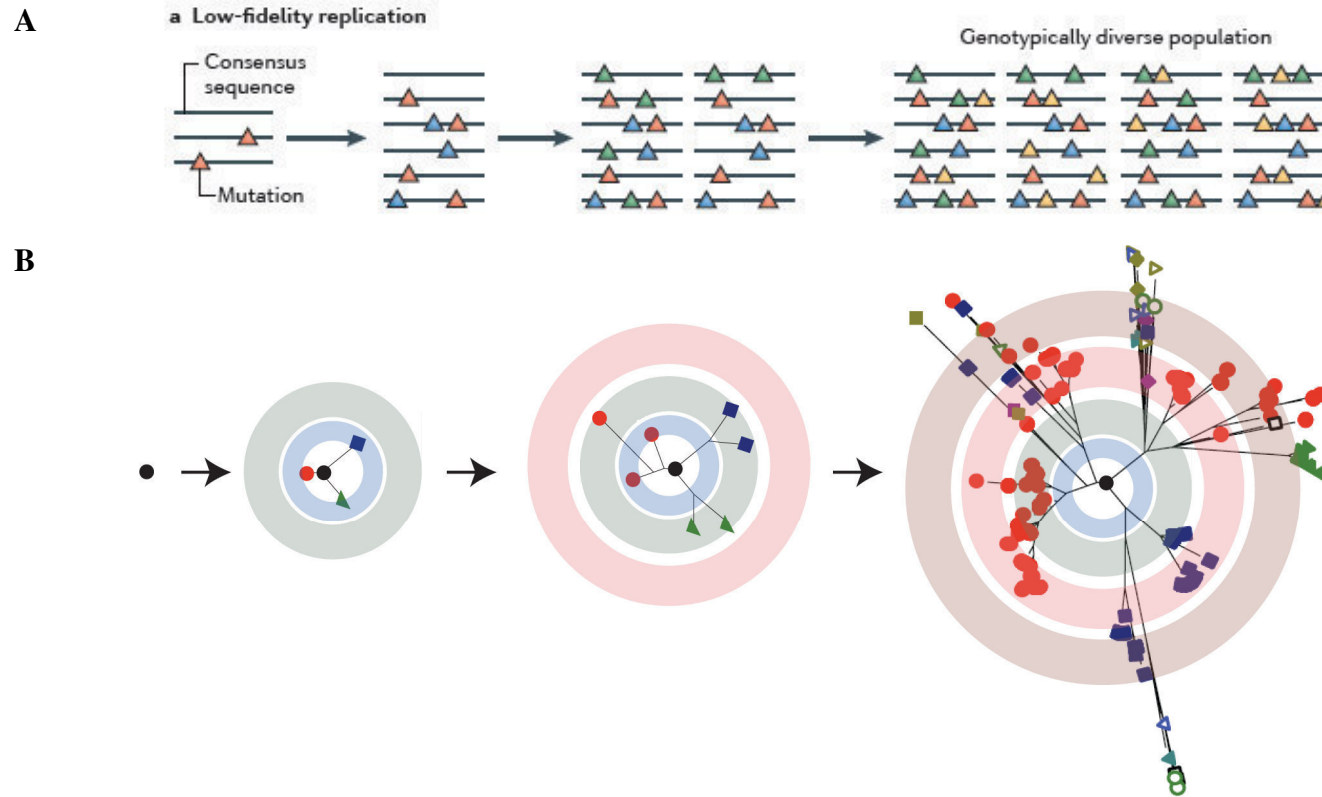
### **1) Les récepteurs de la super famille des immunoglobulines**

Les récepteurs de ce groupe sont des glycoprotéines transmembranaires constituées d'un domaine N-terminal libre, d'une répétition de 2 à 5 domaines de type immunoglobuline et une extrémité C-terminale intra cytoplasmique. Trois molécules de cette famille ont été rapportées en tant que récepteurs pour des entérovirus : PVR (poliovirus receptor), ICAM-1 (inter cellular adhesion molecule 1) et CAR (coxsackievirus adenovirus receptor).

La molécule CD155 ou hPVR (80 kDa) possède 3 domaines de type immunoglobuline (Mendelsohn *et al*, 1989). Elle est phylogénétiquement apparentée aux nectines, est exprimée par les cellules polarisées de la paroi intestinale et les lymphocytes T et intervient dans la reconnaissance et l'adhésion inter cellulaire. Cette molécule sert de récepteur aux trois poliovirus. Le site de liaison poliovirus/CD155 est localisé sur le premier domaine extracellulaire (Koike *et al*, 1991).

La protéine de surface ICAM-1 est constituée de trois domaines extra cellulaires. Sa fonction physiologique, fixer l'intégrine LFA-1 présente sur les leucocytes, permet aux lymphocytes activés de se lier aux cellules endothéliales, ce qui facilite leur transmigration dans les tissus (van de Stolpe et van der Saag, 1996). Cette molécule est exprimée à la surface de nombreux tissus dont l'épithélium nasal. ICAM-1 a été caractérisée en tant que récepteur des rhinovirus (Greve *et al*, 1989) et des entérovirus de l'espèce EV-C : CV-A13, CV-A18 et CV-A21 (Shafren *et al*, 1997a).





**FIGURE 11. Mutations ponctuelles chez les virus à ARN**

**A.** Chaque ligne représente un génome viral. Les mutations acquises lors de chaque cycle de réplication sont représentées par les triangles colorés. Au fur et à mesure des cycles viraux, de nouveaux variants viraux émergent et perdurent selon leur avantage sélectif. (D'après Lauring *et al.*, 2013).

**B.** Un virus avec un taux de mutations élevé (centre) peut générer un répertoire important de mutants (branches) au fur et à mesure des générations (chaque cercle représente un cycle de réplication). Chaque branche indique deux variants reliés entre eux par une mutation ponctuelle.

La protéine CAR sert de récepteur aux adénovirus et aux six coxsackievirus B (Bergelson *et al*, 1997). Il s'agit d'une glycoprotéine transmembranaire de 46 kDa qui compose les jonctions serrées entre les cellules épithéliales intestinales et respiratoires.

## **2) Les récepteurs de la famille des intégrines**

Trois protéines de la famille des intégrines ont été décrites comme récepteurs pour des virus de l'espèce EV-B : les intégrines  $\alpha V\beta 3$ ,  $\alpha V\beta 6$  et  $\alpha 2\beta 1$ . L'intégrine  $\alpha V\beta 3$  ou vitronectine est utilisée par l'échovirus 9 et le CV-A9 (Roivainen *et al*, 1994 ; Ylipaasto *et al*, 2010). L'intégrine  $\alpha V\beta 6$  est exprimée dans plusieurs tissus et sert de récepteur aux génotypes CV-A9 et CV-B1 (Agrez *et al*, 1997 ; Heikkilä *et al*, 2009). L'intégrine  $\alpha 2\beta 1$ , appelée également VLA-2, est un récepteur de l'échovirus 1 (Bergelson *et al*, 1993).

## **3) Le récepteur DAF (Decay Accelerating Factor, CD55)**

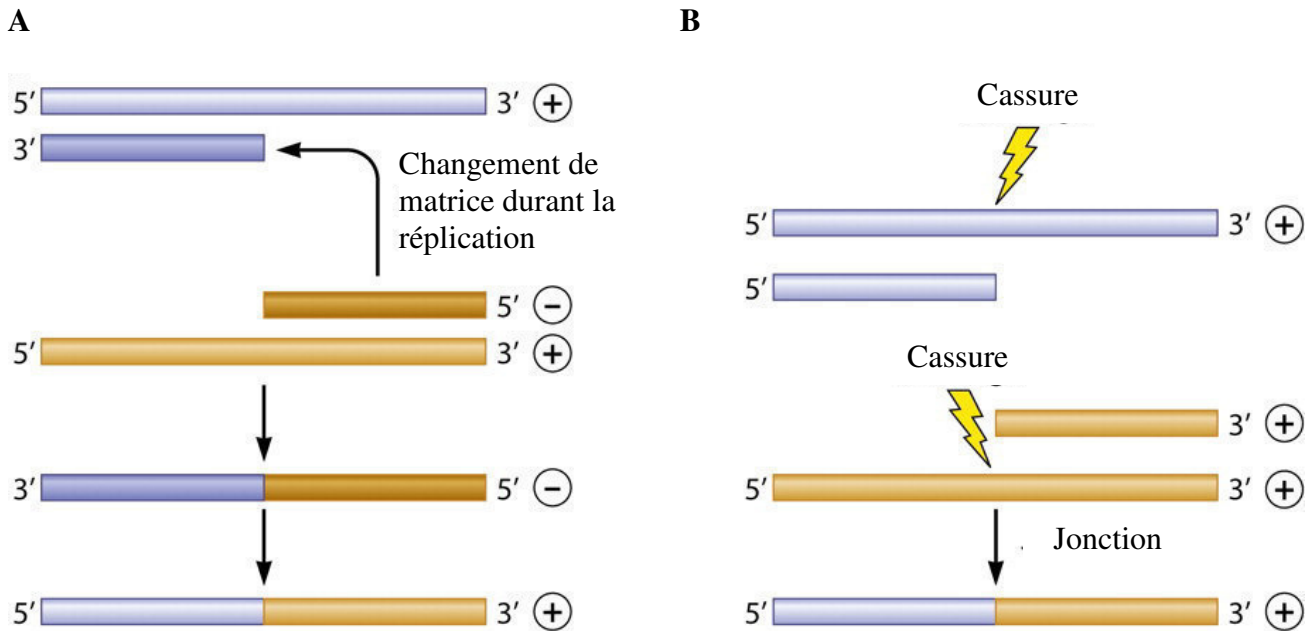
La protéine DAF participe à la régulation de l'activité du complément en inhibant la protéine C3 convertase (Heeger *et al*, 2005). Son rôle de corécepteur a été montré pour plusieurs entérovirus : echovirus 6, 7, 11, 13, 21, 29, 33 ; Coxsackievirus B1, B3, B5, et A21 ; l'entérovirus D70 (Shieh et Bergelson, 2002 ; Alexander et Dimock, 2002 ; Sobo *et al*, 2011). L'interaction des particules virales avec la protéine DAF à la surface cellulaire faciliterait secondairement leur liaison avec le récepteur « principal » (CAR, ICAM-1 ou l'intégrine  $\alpha V\beta 3$  ; Shafren *et al*, 1997b).

# **VI. Réplication virale et diversité génétique des entérovirus**

La grande diversité des entérovirus est la conséquence des événements de mutation et de recombinaisons\* qui surviennent au cours de la réplication des génomes. Ces mécanismes de variation permettent l'adaptation des souches aux diverses pressions de sélection et ils structurent leur évolution et leur dissémination sous la forme de populations\* virales.

## **1) Mutations ponctuelles**

Chez les picornavirus, l'ARN polymérase ARN-dépendante n'effectue pas l'activité correctrice exonucléasique de 3' en 5' (proofreading) présente chez les polymérases



**Figure 12. Recombinaison homologue vs recombinaison non homologue**

A. Lors de la répllication du génome viral, l'ARN polymérase  $3D^{pol}$  débute la synthèse du brin « - » à partir de l'extrémité 3' d'une première matrice ARN brin « + » puis change de matrice ARN à partir de laquelle elle va compléter le brin « - » en cours de synthèse.

B. La recombinaison non homologue consiste en un clivage de deux ARN parentaux suivi d'une ligation des deux fragments.

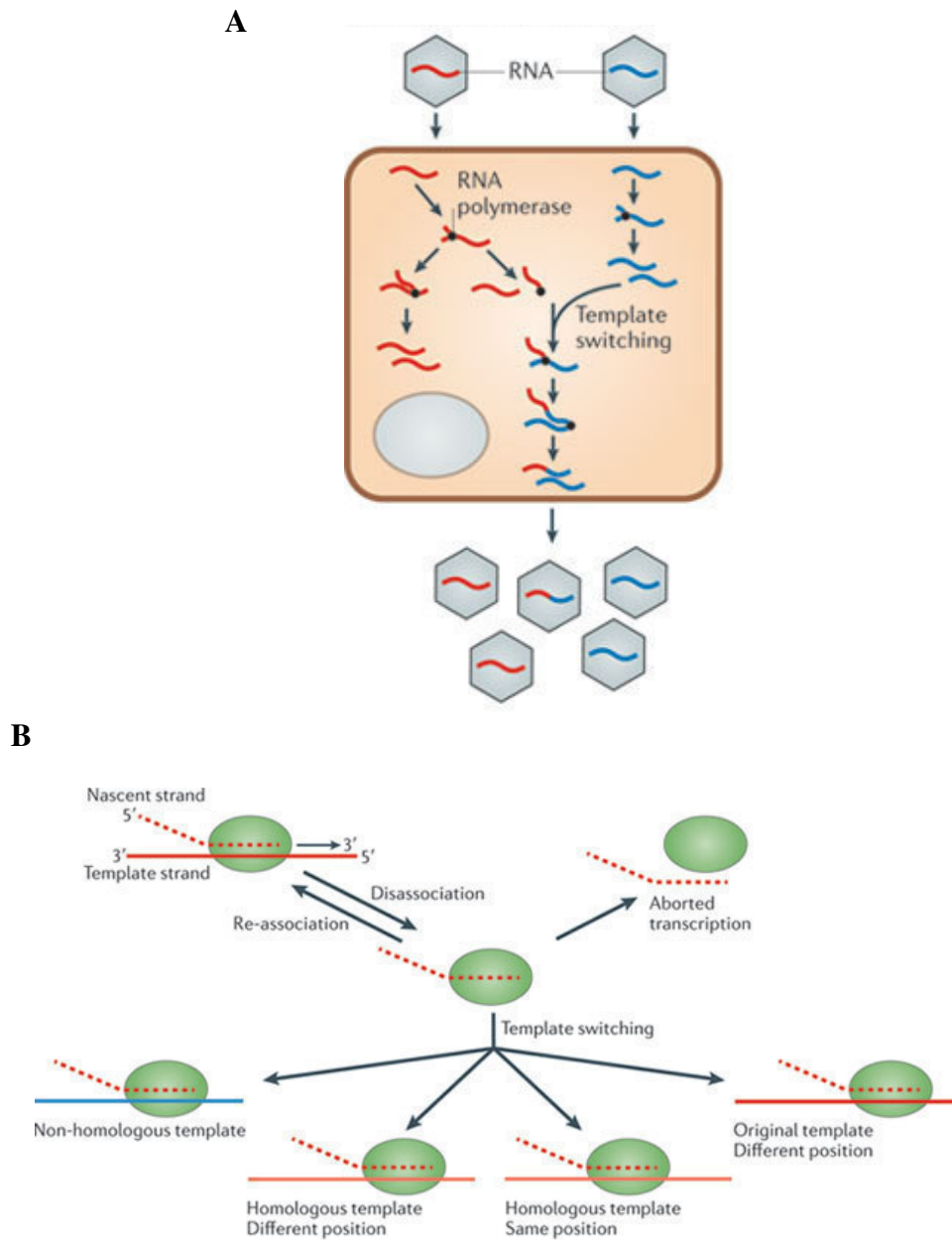
(D'après Domingo et al., 2012).

cellulaires. Au fur et à mesure des cycles viraux, la réplication du génome viral induit des substitutions qui génèrent des mutations ponctuelles dans les génomes néo synthétisés. Le taux moyen d'évolution estimé est compris entre  $10^{-2}$  et  $10^{-4}$  substitutions nucléotidiques par site et par an (Jenkins *et al*, 2002). Sachant que les entérovirus ont un génome d'une longueur de 7500 bases, une mutation aléatoire aura lieu à chaque génome répliqué. La nature de la substitution nucléotidique (transition ou transversion) influence la fidélité de l'ARN polymérase virale des entérovirus : le taux des transitions varie entre  $1.2 \times 10^{-4}$  et  $1.0 \times 10^{-6}$  substitutions/site alors que celui des transversions oscille entre  $3.2 \times 10^{-5}$  et  $4.3 \times 10^{-7}$  substitutions/site (Freistadt *et al*, 2007). De manière générale, le taux de mutations des entérovirus est l'un des plus élevés parmi les virus (Jenkins *et al*, 2002).

La persistance des substitutions est limitée par le processus de sélection naturelle [Figure 11]. Les substitutions qui passent le crible de la sélection dépendent de leur localisation génomique. Les comparaisons des séquences de génomes complets d'entérovirus montrent que le polymorphisme nucléotidique est inégalement réparti. Par exemple, les séquences géniques de la région génomique P1 (protéines structurales) présentent une plus grande capacité à accumuler des mutations. Le polymorphisme est plus élevé en raison d'un plus grand potentiel d'accumulation de mutations dans les boucles protéiques exposées à la surface de la capsid virale. Il en résulte une hétérogénéité de déterminants antigéniques pour le virus, ce qui peut constituer un avantage sélectif relatif entre les populations virales en circulation. Chaque sérotype d'EV est caractérisé par une signature antigénique unique (Brown *et al*, 2003). La survenue des mutations est également soumise aux contraintes imposées par le maintien de structures tridimensionnelles fonctionnelles dans le génome et la capsid virale, et la conservation de sites actifs des enzymes virales.

## **2) Recombinaison génétique**

L'échange de positions génomiques entre deux virus proches parents a été mis en évidence pour la première fois chez le poliovirus en utilisant des marqueurs de résistance à des mutagènes (Ledinko, 1963), ouvrant de nouvelles perspectives en génétique virale. La preuve définitive de la survenue d'évènements de recombinaison a été apportée en montrant que des souches recombinantes de FMDV (Foot and mouth disease virus) contenaient des fragments génomiques des deux virus parentaux (King *et al*, 1982). Le poliovirus a été utilisé comme modèle viral pour comprendre les facteurs moléculaires qui contribuent à la survenue des évènements de recombinaison pendant la réplication du génome viral lorsqu'une cellule est co-infectée par deux souches [Figure 12].



**Figure 13. Recombinaison génétique des entérovirus**

A. Lors d'une co-infection cellulaire par deux virus génétiquement différents (rouge et bleu), un évènement de recombinaison peut se produire (template switching) et générer des virus recombinants.

B. Dès la dissociation de la polymérase virale (vert) et du brin naissant de la matrice virale, la polymérase doit trouver une nouvelle matrice afin de poursuivre la transcription du génome virale. L'évènement de ré-association peut se produire avec la même matrice virale (rouge) à la même position ou à une position différente. Il peut également se réaliser avec une matrice homologoue (orange) ou non homologoue (bleu).

(D'après Simon-Loriere et Holmes, 2011).

### **A. Recombinaison homologue dite « choix de copie »**

La recombinaison homologue virale, aussi appelée « choix de copie » se produit pendant la synthèse du brin de polarité négative (Kirkegaard et Baltimore, 1986). L'ARN polymérase 3D<sup>pol</sup> amorce la synthèse du brin de polarité négative à partir de l'extrémité 3' d'une première molécule d'ARN « + ». Lorsque le complexe de réplication se décroche et poursuit la réplication du brin « - » sur une autre matrice ARN « + », une molécule d'ARN mosaïque est générée [Figures 12A et 13]. Ce modèle moléculaire propose que le complexe réplcatif se détache de la matrice A entraînant le brin naissant incomplet dont l'extrémité 3' s'hybride à la matrice B et le complexe réplcatif poursuit son élongation (Pilipenko *et al*, 1995). Le décrochage du complexe de réplication pourrait être causé par la présence de structures secondaires dans la matrice initiale. La liaison du complexe de réplication à un autre génome à proximité est liée à la similitude nucléotidique entre les deux matrices parentales (Tolskaya *et al*, 1987). Le degré de parenté génétique est un facteur majeur de la probabilité de la recombinaison virale (Kirkegaard et Baltimore, 1986) : les événements de recombinaison intratypique (entre des souches de même sérotype) ont 100 fois plus de chance de se produire par rapport aux événements de recombinaison intertypique (souches de sérotypes différents).

### **B. Recombinaison non homologue « non réplcative »**

Des virus viables ont été obtenus *in vitro* après co-transfection dans une cellule hôte de deux ARN génomiques viraux dont la traduction et la réplication virale ont été inactivées (Gmyl *et al*, 1999). Les génomes viraux néoformés présentent des recombinaisons entre des segments non homologues, ce qui suggère l'existence d'un mécanisme non réplcatif de recombinaison virale [Figure 12B] (Gmyl *et al*, 1999). Le mécanisme moléculaire proposé pour expliquer cette recombinaison est une liaison par trans-estérification entre les matrices virales non chevauchantes (Gmyl *et al*, 2003). La liaison ne dépend pas de la formation d'un hétéroduplex entre les matrices génomiques mais nécessite la présence de deux groupements chimiques distincts sur chacune d'elles : un groupement 3'-phosphate et un groupement 5'-hydroxyl. La jonction s'effectue préférentiellement au niveau d'une région simple brin et peu structurée du génome viral ; la proportion de chaque matrice parentale influence l'efficacité de la recombinaison non homologue (Austermann-Busch et Becher, 2012).

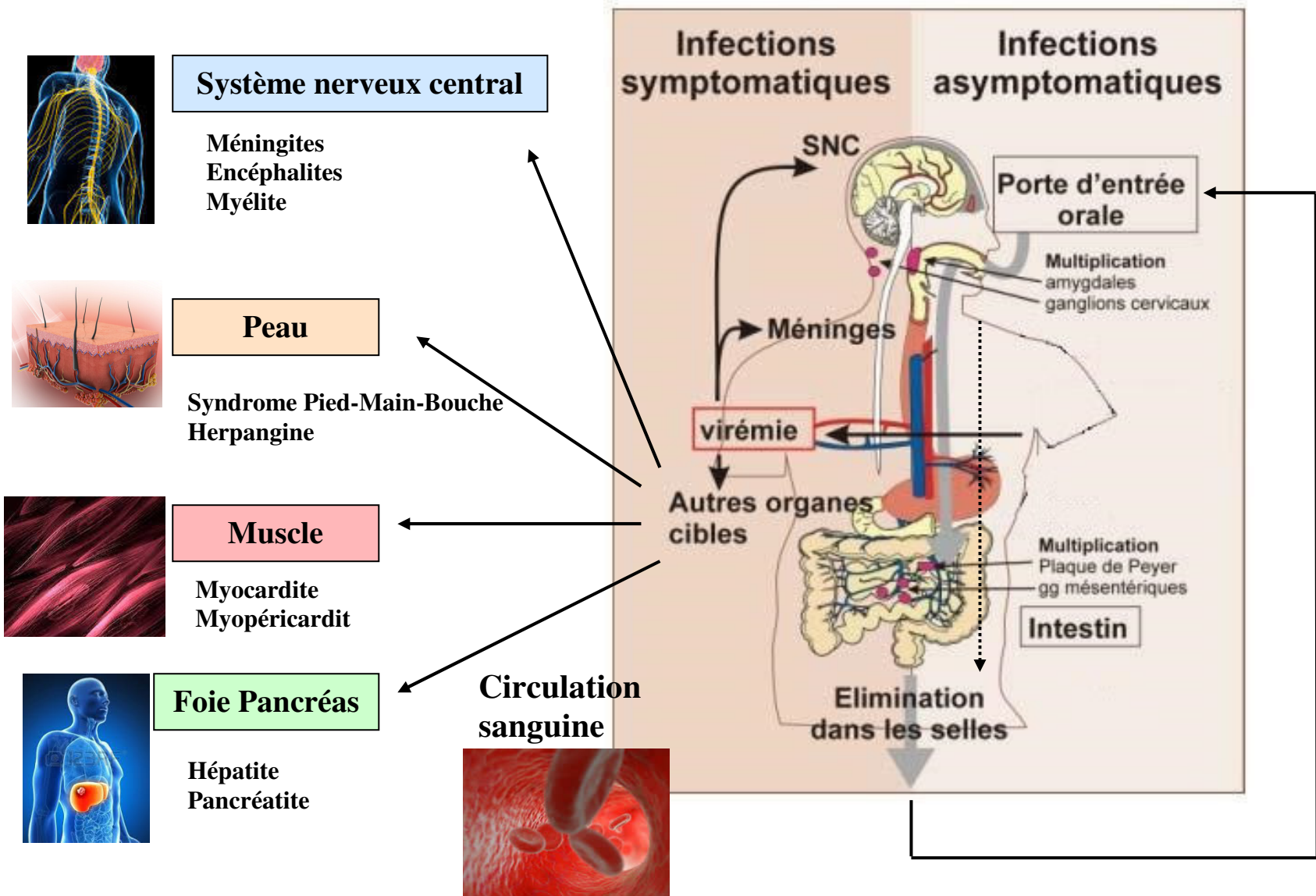


FIGURE 14. Physiopathologie des entérovirus

Modifié d'après <http://www.microbes-edu.org/etudiant/imgpicor/14.jpg>

## VII. Principales pathologies et diagnostic des entérovirus

Les infections à EV symptomatiques sont liées une grande diversité de présentations cliniques. Après ingestion ou inhalation, le virus se multiplie dans des sites primaires : les muqueuses (*lamina propria*) pharyngiennes (amygdales) et intestinales (plaques de Peyer). Chez la majorité des individus, les réponses immunitaires initiées dès ce stade contiennent l'infection virale qui demeure peu ou pas symptomatique. Secondairement, la virémie (virus disséminé dans la circulation sanguine) peut être à l'origine d'un syndrome dont la sévérité dépend du statut immunitaire, de la souche, etc... [Figure 14].

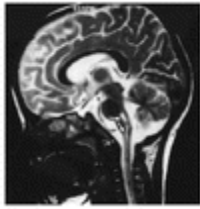
### 1) Atteintes neuroméningées

De nombreux sérotypes d'EV peuvent infecter le SNC. Le poliovirus (espèce EV-C) est le sérotype neurotrope le plus connu en raison des nombreuses épidémies de paralysies flasques occasionnées dans le passé. Parmi les autres EV, l'échovirus 30 (EV-B) et l'EV-A71 (EV-A) sont particulièrement neurotropes et épidémiques (Pallansch et Roos, 2007). Les atteintes du SNC par les EV se manifestent sous la forme de trois entités cliniques principales : les méningites aiguës (dites aseptiques), les encéphalites et les myélites. Les méningites aseptiques sont des inflammations causées par beaucoup d'entérovirus et particulièrement ceux de l'espèce EV-B : echovirus et coxsackievirus B (Pallansch et Roos, 2007). La réaction inflammatoire qui se développe dans les méninges provoque des symptômes généraux (céphalées, fièvre, somnolence, raideur de nuque, photosensibilité, nausées, vomissements...). Le virus est retrouvé très rapidement après le début des symptômes dans le liquide céphalo-rachidien (LCR), la charge virale dans le LCR dépend de l'âge des patients et du sérotype de l'entérovirus en cause (Volle *et al*, 2014). Les méningites à entérovirus sont spontanément résolutes et moins sévères que les méningites d'origine bactérienne, ce qui indique que la multiplication virale reste cantonnée à la périphérie du cerveau ; les sites tissulaires précis dans lesquels les entérovirus se multiplient ne sont pas connus.

### 2) Autres pathologies

Les entérovirus peuvent atteindre de nombreux autres tissus tels que la peau, les yeux, le cœur, le foie, le pancréas ou les muscles. Les infections de la peau et des muqueuses provoquent le syndrome pied-main-bouche et l'herpangine qui sont détaillés dans la 3<sup>ème</sup> partie de cette introduction. Les infections à entérovirus EV-D70 et CV-A24 peuvent causer





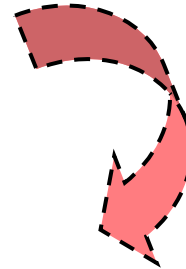
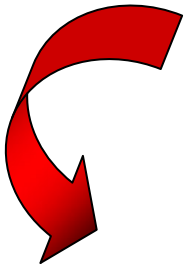
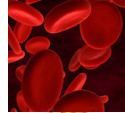
## Signes cliniques

- Fièvre
- Syndromes PMB
- Herpangine
- Symptômes neurologiques
- Autres



## Prélèvements biologiques

- Gorge
- Vésicules
- Selles
- LCR
- Sang



## Biologie moléculaire

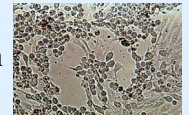
### Détection rapide par PCR en temps réel

- Extraction des acides nucléiques
- RT-qPCR – Région 5'NC



## Isolement - culture cellulaire

- Culture cellulaire (RD, Vero, MRC5, Hela)
- Inoculation des prélèvements
- Effet cytopathogène ou non
- Immunofluorescence



## Génotypage

- RT-PCR – Locus 1D<sup>VP1</sup>
- Séquençage

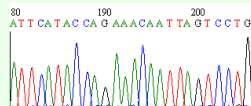
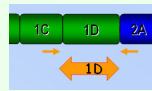


Figure 15. Diagnostic des entérovirus causant un syndrome PMB

des conjonctives hémorragiques aiguës, caractérisées par une incubation de courte durée. Les symptômes sont une augmentation excessive du larmoiement, des douleurs oculaires, un œdème périorbitaire et une rougeur de la conjonctive (Yin-Murphy, 1984). Les EV peuvent être à l'origine de myocardites (Andréoletti *et al*, 2009), de symptômes respiratoires (Pallansch et Roos, 2007). De nombreux travaux suggèrent un lien entre les infections à EV et la survenue du diabète insulino-dépendant mais les mécanismes pathogéniques ne sont pas établis (Kondrashova *et al*, 2014).

### **3) Diagnostic**

Le diagnostic des infections neuro-méningées à entérovirus a pour but d'identifier l'agent infectieux et d'exclure une infection bactérienne ou herpétique ; il consiste à détecter le génome viral dans un prélèvement de LCR. Le diagnostic des atteintes neuro-méningées est effectué avec le prélèvement de liquide céphalorachidien (LCR) obtenu par une ponction lombaire généralement réalisée lors de l'admission des patients. La présence du génome viral dans ce liquide biologique normalement stérile indique un passage du virus à travers les barrières sang/LCR (hémato-méningées) et/ou sang/parenchyme cérébral (hémato-encéphalique).

Le diagnostic de la maladie pied-main-bouche est réalisé avec plusieurs prélèvements : gorge, vésicules cutanées ou buccales et selles [Figure 15]. Ces prélèvements permettent la détection et l'isolement de la souche de virus (Ooi *et al*, 2007), contrairement au LCR à partir duquel le taux d'isolement viral est faible. Les isolats obtenus à partir de vésicules présentent un intérêt particulier puisqu'ils sont caractéristiques de l'infection systémique.

### **4) Identification moléculaire (génotypage)**

L'identification de la souche par des techniques moléculaires permet d'apporter des informations épidémiologiques sur le sérotype en cause et elle est effectuée en utilisant le prélèvement clinique utilisé pour le diagnostic virologique. L'avantage principal est la rapidité d'obtention du résultat. Un autre avantage est la possibilité d'intégrer le génotypage (y compris l'étape de séquençage) dans le flux général des techniques moléculaires réalisées dans un laboratoire de diagnostic (RT-PCR et séquençage de la région génomique 1D<sup>VP1</sup> ou VP4/VP2).

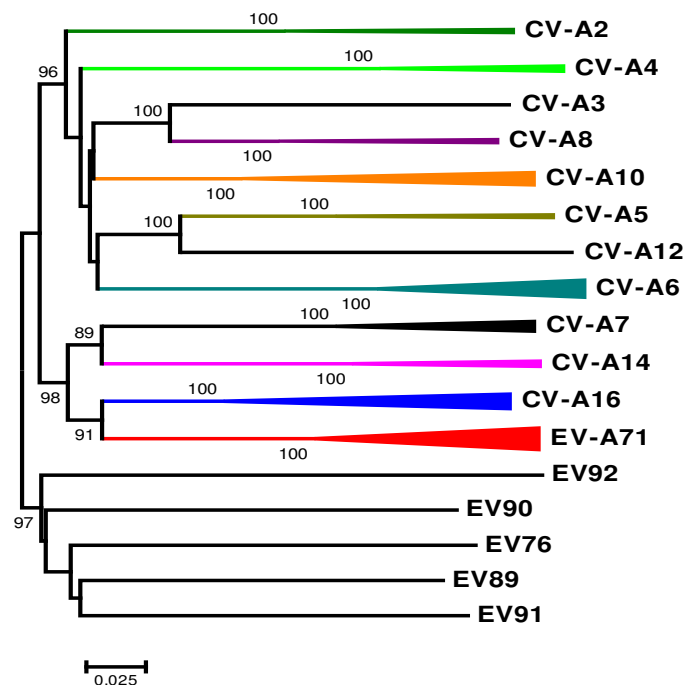
L'identification peut aussi être effectuée en utilisant la souche de virus isolée en culture cellulaire. L'isolement viral consiste à inoculer le virus contenu dans un prélèvement

clinique à une lignée cellulaire permissive pour le virus. Il existe un grand éventail de lignées cellulaires : par exemple, les cellules humaines de rhabdomyosarcome (RD) et les cellules rénales de singe vert d'Afrique (Vero) sont couramment utilisées pour l'isolement des souches d'EV-A (Schmidt *et al*, 1975). D'autres types cellulaires sont couramment utilisés comme les cellules embryonnaires fibroblastiques de poumon humain à durée de vie limitée (cellules MRC-5, Medical Research Council 5).

**L'ENTEROVIRUS TYPE  
A71 ET LES  
SEROTYPES  
APPARENTES DE  
L'ESPECE A**

**Tableau 3. Sérotypes d'EV-A reconnus**

Sérotipe	Souche prototype	Origine et année d'isolement	Reference	Pathologies associées
CV-A2	Fleetwood	USA - 1947	Sickles and Dalldorf, 1949	Herpangine et syndrome PMB
CV-A3	Olson	USA - 1948	Sickles and Dalldorf, 1949	Herpangine et syndrome PMB
CV-A4	High Point	USA - 1950	Melnick, 1950	Herpangine et syndrome PMB
CV-A5	Swartz	USA - 1950	Dalldorf and Sickles, 1953	Herpangine et syndrome PMB
CV-A6	Gdula	USA - 1949	Dalldorf <i>et al.</i> , 1949	Herpangine et syndrome PMB
CV-A7	Parker	USA - 1949	Dalldorf <i>et al.</i> , 1949	Rarement isolé
CV-A8	Donovan	USA - 1949	Dalldorf <i>et al.</i> , 1949	Herpangine et syndrome PMB
CV-A10	Kowalik	USA - 1950	Dalldorf <i>et al.</i> , 1949	Herpangine et syndrome PMB
CV-A12	Texas-12	USA - 1948	Melnick, 1950	Rarement isolé
CV-A14	G-14	Afrique du Sud - 1950	Gear, 1958	Rarement isolé
CV-A16	G-10	Afrique du Sud - 1951	Gear, 1958	Herpangine et syndrome PMB
EV-A71	BrCr	USA - 1970	Schmidt <i>et al.</i> , 1974	Syndrome PMB et complications neurologiques/pulmonaires
EV-A76	FRA91-10369	France - 1991	Oberste <i>et al.</i> , 2005	Rarement isolé
EV-A89	BAN00-10359	Bangladesh - 2000	Oberste <i>et al.</i> , 2005	Rarement isolé
EV-A90	BAN99-10399	Bangladesh - 1999	Oberste <i>et al.</i> , 2005	Rarement isolé
EV-A91	BAN00-10406	Bangladesh - 2000	Oberste <i>et al.</i> , 2005	Rarement isolé
EV-A92	USA/GA99-RJg-7			Isolé chez le singe
EV-A114	BAN-11610			Rarement isolé
EV-A119	09C13CMR			Rarement isolé
EV-A120	184977-PNG-09			Rarement isolé
EV-A121	V13-0682 (India)			Rarement isolé
SV19	M19s			Isolé chez le singe
SV43	OM112t			Isolé chez le singe
SV46	OM22			Isolé chez le singe
BA13	A13			Isolé chez le singe



**Figure 16. Diversité de sérotypes au sein des EV-A**

Arbre phylogénétique construit par la méthode du Neighbor-Joining à partir d'un échantillon de séquences P1 d'EV-A. Chaque sérotipe est représenté par une couleur différente. La fiabilité des nœuds est exprimée en pourcentage de bootstrap (1000 itérations).

L'espèce EV-A se compose de 21 sérotypes d'entérovirus humains (CV-A2, CV-A3, CV-A4, CV-A5, CV-A6, CV-A7, CV-A8, CV-A10, CV-A12, CV-A14, CV-A16, EV-A71, EV-A76, EV-A89, EV-A90, EV-A91, EV-A92, EV-A114, EV-A119, EV-A120 et EV-A121) et de 4 sérotypes d'entérovirus simiens (SV19, SV43, SV46 et BA13) [Table 3].

Les sérotypes humains constituent trois groupes monophylétiques [Figure 16] :

- L'un comprend les types principalement associés à l'herpangine (CV-A2, CV-A3, CV-A4, CV-A5, CV-A6, CV-A7, CV-A8, CV-A10, CV-A12)
- Le deuxième, les types associés à la maladie pied-main-bouche (CV-A14, CV-A16 et EV-A71)
- Le troisième groupe intègre des types identifiés récemment et plus rarement isolés (EV-A76, EV-A89, EV-A90...)

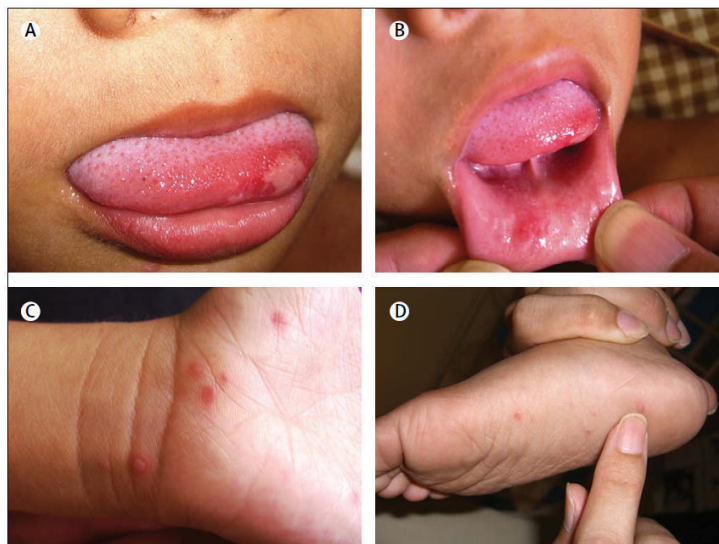
Les maladies les plus fréquentes chez l'homme causées par les sérotypes de l'espèce A sont des manifestations inflammatoires cutanéomuqueuses : la maladie pied main bouche et l'herpangine. Ces maladies contagieuses sont épidémiques et affectent principalement les jeunes enfants de moins de 5 ans ; ces formes symptomatiques sont majoritairement bénignes.

Les complications les plus sévères de la maladie pied-main-bouche sont observées pendant l'infection par des souches EV-A71 ; elles sont liées au neurotropisme de cet entérovirus. Les complications neurologiques affectent les très jeunes enfants et se présentent sous la forme d'encéphalite du tronc cérébral, de paralysie flasque (type poliomyélite) et d'atteintes cardiopulmonaires qui nécessitent une prise en charge en soins intensifs.

## **I. Manifestations cliniques**

### **1) Maladie pied-main-bouche**

La première caractérisation formelle de la maladie pied-main-bouche a été rapportée en 1960 à Birmingham en Angleterre (Alsop *et al*, 1960). Les signes cliniques de la maladie pied-main-bouche se déclarent 3 à 7 jours après un contact avec une personne infectée. Ils débutent par une fièvre modérée, une diarrhée, des arthralgies et un malaise général. Par la suite, un exanthème vésiculaire cutané apparaît principalement sur la plante des pieds, la paume des mains et la muqueuse de la cavité buccale (langue, palais et gencives) [Figure 17] (Alsop *et al*, 1960). Ces vésicules cutanées ne provoquent pas de démangeaisons mais sont douloureuses. La maladie PMB peut aussi se manifester sous la forme d'éruptions cutanées



**Figure 17. Exanthème vésiculaire cutanée du syndrome pied-main-bouche**  
Ulcères sur la langue (A), les lèvres (B) ; lésions maculaires et vésiculaires sur le poignet (C) et le pied (D) observées chez des enfants infectés par des EV-A.  
[Ooi *et al.*, 2010]



**Figure 18. Eczéma coxsackium : forme atypique du syndrome pied-main-bouche**  
Lésions cutanées sur l'intégralité du corps  
[Mathes *et al.*, 2013]

maculopapuleuses (sans vésicules) sur les fesses, les genoux et les coudes. Plus rarement, une atteinte du système respiratoire supérieur, une gastroentérite, une bronchiolite ou une pneumonie peuvent être retrouvées lors de l'infection (Ooi *et al*, 2010). Récemment, des formes présentées comme atypiques de maladie pied-main-bouche ont été rapportées : eczéma coxsackium, lésions cutanées sur l'intégralité du corps [Figure 18] (Mathes *et al*, 2013).

## **2) Herpangine**

L'herpangine est un syndrome marqué par l'apparition d'une fièvre accompagnée de mal de gorge et d'un érythème douloureux dans la cavité buccale postérieure comprenant les plis du pharynx antérieur, la luette, les amygdales et le palais mou (Zahorsky, 1951). Chez certains enfants, l'infection se propage à d'autres parties de la bouche, comme les gencives et la langue, de sorte que la distinction des signes cliniques avec ceux de la maladie PMB est parfois difficile.

## **3) Paralysie flasque de type poliomyélite**

Les paralysies flasques ressemblant à la poliomyélite peuvent être des symptômes d'atteintes neurologiques causées par de nombreux entérovirus (Melnick, 1984). Elles se caractérisent par une apparition soudaine de faiblesses musculaires au niveau des membres inférieurs dans la plupart des cas. Ces signes cliniques surviennent à cause de l'infection et la destruction des motoneurons dans la corne antérieure de la moelle épinière, un processus pathologique proche de l'infection par les souches de poliovirus (McMinn, 2002). Contrairement aux paralysies induites par le poliovirus, les paralysies causées par l'EV-A71 ne perdurent pas au-delà de 60 jours et présentant un taux de rétablissement plus élevé par rapport aux infections par des poliovirus.

## **4) Encéphalite du tronc cérébral**

L'atteinte du tronc cérébral est une encéphalite et l'une des plus sévères manifestations neurologiques de l'infection à EV-A71 (Ooi *et al*, 2010). Elle peut conduire au décès de l'individu infecté. Elle est classée en trois stades de sévérité : (i) encéphalite sans complications, (ii) dérégulation du système nerveux autonome et (iii) œdème pulmonaire (Wang *et al*, 2008). Certains symptômes précoces (contractions myocloniques, ataxie, tremblements, nystagmus (mouvement d'oscillation involontaire du globe oculaire) ou paralysies oculaires et bulbaires) indiquent une atteinte du tronc cérébral (stade n°1).

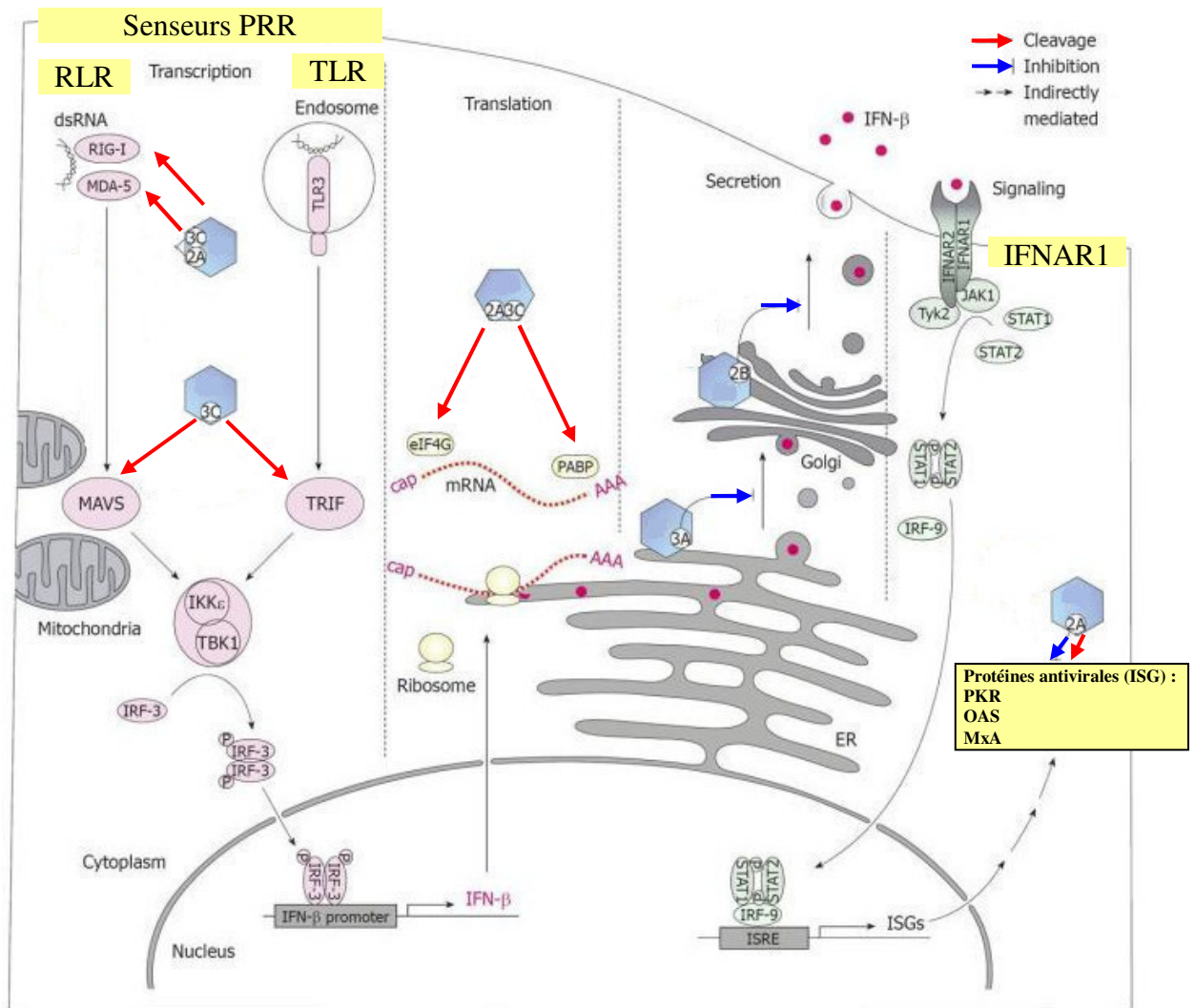


L'individu infecté peut développer rapidement des symptômes tels qu'une tachycardie, une tachypnée et une hypertension, signes d'une dérégulation du système nerveux autonome (stade n°2). L'hypertension associée à cette infection virale est provoquée par une atteinte des centres respiratoires et vasomoteurs au niveau de la médulla. L'infection virale peut atteindre le troisième stade de sévérité, caractérisé par certains symptômes comme une hyperglycémie, une cyanose ou/et un syndrome lié à un œdème pulmonaire et des troubles cardiaques. Ces atteintes respiratoires peuvent conduire au décès de l'individu infecté dans 80 à 90 % des cas (Lum *et al*, 1998 ; Ooi *et al*, 2010). Le processus physiopathologique conduisant à la survenue de l'œdème pulmonaire sévère n'est pas clairement compris à l'heure actuelle. L'hypothèse retenue est la combinaison d'une réponse inflammatoire systémique et d'une inflammation de la substance grise de la *medulla oblongata*. Cette réponse inflammatoire serait liée à la libération de cytokines et chemokines (IL-6, IL-10, IL-13 et TNF $\alpha$ ) entraînant une augmentation de la perméabilité vasculaire pulmonaire. L'infection de la *medulla oblongata* aboutirait à une dérégulation du système autonome, favorisant l'œdème pulmonaire sévère. Ce type d'atteinte peut provoquer des séquelles neurologiques pour le reste de la vie des individus (Tsai *et al*, 2014).

## II. Pouvoir pathogène

Les EV-A responsables de la maladie pied-main-bouche peuvent se transmettre *via* le contact avec des sécrétions orales, des fluides vésiculaires ou des gouttelettes respiratoires contenant le virus. Comme les autres entérovirus, les sites initiaux de la multiplication virale sont l'oropharynx (amygdales) et l'intestin (plaques de Peyer). Le virus peut être disséminé par le sang et le système lymphatique vers les ganglions lymphatiques régionaux et d'autres tissus (peau, muqueuses, méninges, bulbe rachidien, moelle épinière, foie, pancréas...). La progression du virus vers ces différents tissus dépend du tropisme de la souche et des réponses immunes de l'hôte.

Plusieurs réponses immunitaires se mettent en place lors d'une infection virale : libération de cytokines, production d'anticorps et activation de lymphocytes T cytotoxiques. La réponse immune innée intracellulaire génère un environnement défavorable pour le cycle viral et favorise la présentation d'antigènes viraux. Cette réponse antivirale est activée grâce à des senseurs moléculaires appelées « Pattern Recognition Receptors » (PRR). Certains d'entre eux reconnaissent des motifs conservés chez les agents pathogènes (ou PAMP) tels les ARN double brin (Kato *et al*, 2006). Trois familles de récepteurs cytoplasmiques ont été décrites :



**Figure 19. Mécanisme moléculaire d'inhibition de la synthèse d'interféron lors d'une infection par un entérovirus**

L'infection d'une cellule hôte par un entérovirus enclenche une réponse immunitaire antivirale, s'amorçant par la reconnaissance de l'ARN double brin viral par les senseurs moléculaires, RLR et TLR. Ces derniers vont activer des cascades cellulaires aboutissant à la synthèse d'IFN- $\beta$ . La fixation de cette cytokine à ses récepteurs, IFNAR1, enclenche la synthèse de protéines ayant des actions antivirales telles que PKR, OAS et MxA. Le virus peut limiter cette réponse antivirale par l'action de ses deux protéases 2A<sup>pro</sup> et 3C<sup>pro</sup>, ayant la capacité de cliver de nombreux intermédiaires de la cascade cellulaire de l'IFN- $\beta$  ou d'inhiber leurs synthèses (losange bleu).

Légende : IFN: Interféron; IRF: Interferon regulatory factor; ISGs: Interferon stimulated genes; MDA: Melanoma differentiation-associated gene; dsRNA: Double-stranded RNA; PKR: Protein kinase R; TLR: Toll-like receptor; MAVS: Mitochondrial antiviral signaling protein; ISRE: Interferon stimulated response element; OAS: Oligoadenylate synthetase; RIG-I: Retinoic acid-inducible gene I.

(Modifié d'après Dotzauer et Kraemer, 2012)

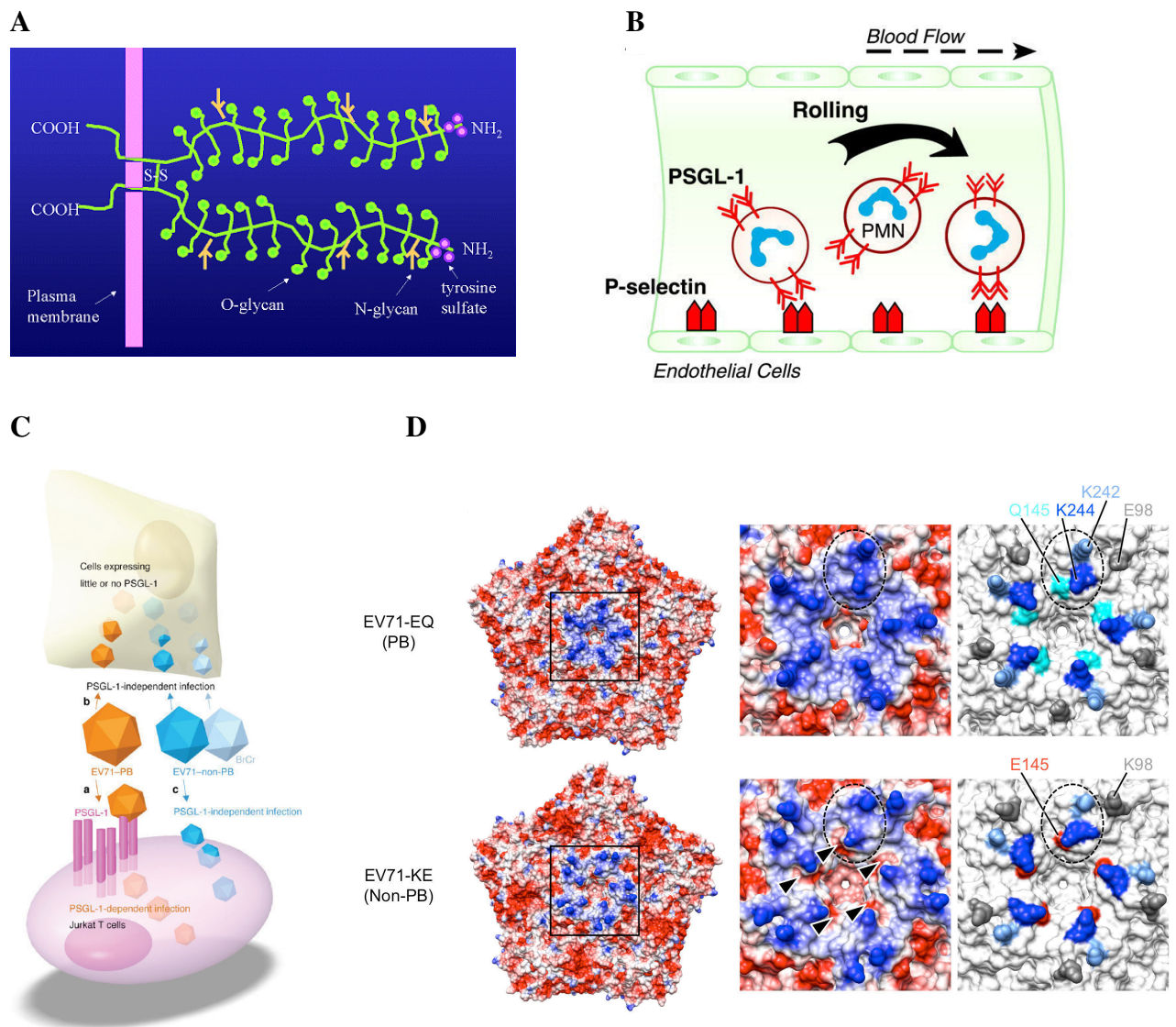
les récepteurs de type RIG-I (RLR), de type Toll (TLR) et de type NOD (NLR). L'activation de ces récepteurs enclenche une suite de réactions impliquant l'activation de la protéine mitochondriale de signalisation antivirale (MAVS, Dotzauer et Kraemer, 2012). Celle-ci favorise la production de cytokines pro-inflammatoires (voie cellulaire NF- $\kappa$ B) et de protéines antivirales (PKR, OAS, MxA) *via* la voie de l'INF- $\beta$  [Figure 19].

Des processus d'échappement qui atténuent les effets de la réponse immune innée antivirale se sont mis en place au cours de l'évolution des entérovirus. La protéase virale 3C<sup>pro</sup> du poliovirus induit un arrêt de la grande majorité des synthèses protéiques cellulaires dont les cytokines et les senseurs PRR (Feng *et al*, 2014). Par ailleurs, la protéase 2A<sup>pro</sup> de l'EV-A71 peut bloquer la signalisation induite par l'interféron  $\beta$  en limitant la production de cette cytokine ou celle des récepteurs à l'INF appelés IFNAR1 (Lu *et al*, 2012 ; Wang *et al*, 2013).

La réponse immune adaptative prend le relais et limite la dissémination du virus dans l'organisme hôte et la réponse humorale joue un rôle prédominant dans la résolution spontanée de l'infection par un entérovirus. L'expérience de plusieurs décennies de vaccination orale contre les poliovirus montre que la présence d'une mémoire immunitaire produisant une immunité des muqueuses permet de limiter rapidement les infections virales secondaires et la diffusion interindividuelle de ces virus (Ogra, 1984). La réponse à médiation cellulaire est aussi mise en place au cours de l'infection à EV-A71 (Dang *et al*, 2014). Cette réponse immune spécifique peut également être contournée. Par exemple, il a été montré *in vitro* que l'EV-A71 peut infecter de nombreuses populations de cellules immunitaires telles que la lignée Jurkat (Lymphocytes T), la lignée de macrophages THP-1, les cellules mononuclées du sang périphérique et les cellules dendritiques humaines (Chen et Yeh, 2009). L'infection des cellules immunitaires par ce virus enclenche la production de cytokines pro inflammatoires (TNF $\alpha$  ou MIF, facteur inhibiteur de migration de macrophages), ce qui pourrait jouer un rôle dans les infections sévères.

### **III. Récepteurs cellulaires de l'EV-A71 et des types apparentés**

Plusieurs molécules de la surface cellulaire ont été décrites comme récepteurs utilisés par l'EV-A71 et plusieurs autres sérotypes de l'espèce EV-A : il s'agit de PSGL-1, SCARB2, l'annexine II, de l'acide sialique et de la lectine DC-SIGN [Tableau 2].



**Figure 20. PSGL-1 : protéine membranaire des leucocytes et récepteur d'attachement de certaines souches d'EV-A71**

(A) Structure du récepteur PSGL-1. Il s'agit d'un polypeptide transmembranaire riche en résidus proline, sérine et thréonine. Les glycanes O-linked portés par PSGL-1 peuvent subir deux types de modifications post-traductionnelles lui permettant d'être un contre récepteur de la sélectine P, une fucosylation (1,3) et une sialylation (2,3 ; <http://bme.virginia.edu/ley/psgl-1.html>).

(B) Rôle physiologique de PSGL-1 chez les leucocytes. La fixation de la Sélectine P au récepteur PSGL-1 permet le rolling des leucocytes vers les cellules endothéliales en conditions inflammatoires (**modifié d'après Zimmerman, 2001**).

(C) PSGL-1 : récepteur de certains EV-A71. La protéine membranaire PSGL-1, exprimée chez les leucocytes est un récepteur fonctionnel pour certaines souches d'EV-A71 (EV71-PB). Les souches EV71-non-PB infectent les leucocytes sans utiliser le récepteur PSGL-1. L'ensemble des souches d'EV-A71 infectent d'autres types cellulaires qui n'expriment pas ou peu le récepteur PSGL-1 (**d'après Nishimura et al., 2009**).

(D) Comparaison des structures de capsides des souches EV-71-PB et EV-71-non-PB. Surface électrostatique des capsides des EV-A71 pouvant se lier ou non au récepteur PSGL-1. Des différences de structures s'observent notamment au niveau de l'axe d'ordre 5 d'un protomère (**d'après Nishimura et al., 2013**).

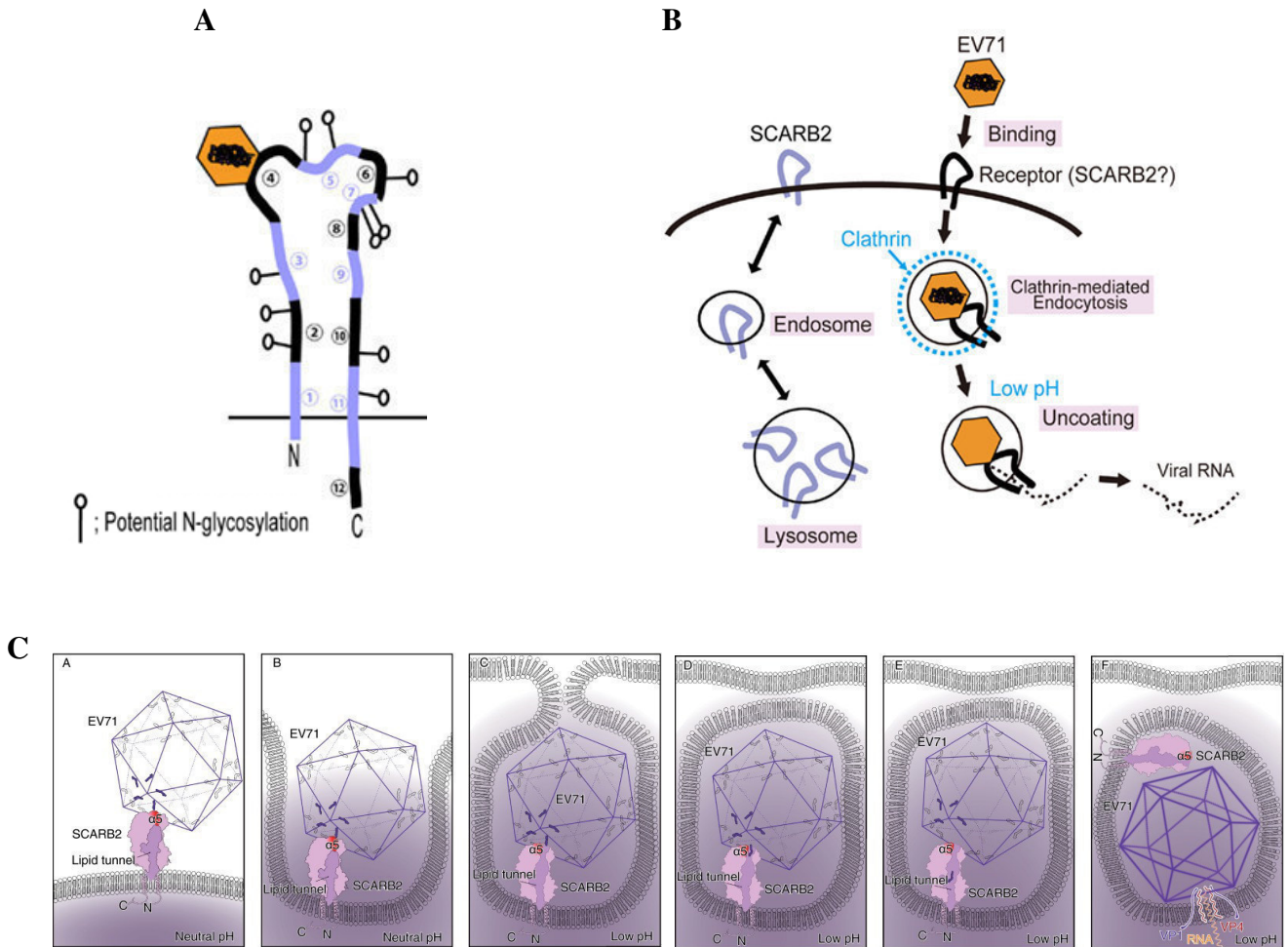
## **1) PSGL-1 (P-selectin glycoprotein ligand-1)**

Les atteintes sévères consécutives à l'infection par l'EV-A71 (œdème cardiopulmonaire) se caractérisent par un taux élevé de cytokines pro-inflammatoires et une diminution du nombre de lymphocytes T. Un lien entre la lymphopénie et l'infection à EV-A71 a été recherché et PSGL-1 a été identifiée comme un récepteur fonctionnel (Nishimura *et al*, 2009). Il s'agit d'une sialomucine exprimée à la surface des leucocytes des lignées myéloïde et lymphoïde [Figure 20a]. Cette protéine est impliquée dans l'attachement et le « rolling » des leucocytes effecteurs sur les cellules endothéliales activées, pendant la réponse inflammatoire [Figure 20b].

PSGL-1 n'est pas le récepteur cellulaire de toutes les souches d'EV-A71 et plusieurs facteurs interviennent dans le mécanisme d'adsorption [Figure 20c]. Pour qu'une souche d'EV-A71 puisse se lier à PSGL-1, celle-ci doit avoir subi des modifications post traductionnelles : sulfatation des tyrosines présentes dans la région N-terminale de la protéine (Nishimura *et al*, 2010). La liaison entre une souche d'EV-A71 et le récepteur PSGL-1 tient aussi à la présence d'une combinaison de deux acides aminés dans la protéine de capsid VP1 [Figure 20d] : une glycine (G) ou une glutamine (Q) à la position 145 combinée à un acide aspartique (E) en position 98 (Nishimura *et al*, 2013). La lignée cellulaire murine L929 a été rendue sensible à l'infection par l'EV-A71 par l'expression stable la protéine humaine PGSL-1 (L-PSGL-1.1). Les cellules transfectées ont été utilisées pour analyser le rôle de cette protéine dans la sensibilité à l'infection virale. Les souches d'EV-A71 qui se lient à PSGL1 se multiplient dans cette lignée cellulaire mais à des taux faibles, ce qui suggère la nécessité de mutations secondaires pour une multiplication virale optimale (Miyamura *et al*, 2011). Le récepteur PSGL-1 ne semble pas être en mesure d'induire l'internalisation et la décapsidation des particules virales, et serait simplement une molécule d'attachement des virions (Yamayoshi *et al*, 2013). La multiplication de différentes souches d'autres types d'EV-A a aussi été comparée dans les cellules L-PSGL-1.1 (Nishimura et Shimizu, 2012). L'étude indique que les types CV-A3, CV-A10, CV-A12, CV-A14 et CV-A16 infectent les cellules mais le récepteur PSGL-1 n'est pas requis.

## **2) SCARB2 (human scavenger receptor class B, member 2)**

L'origine de la sensibilité des cellules RD (rhabdomyosarcome) à l'infection par les souches d'EV-A71 a été recherchée par transfection d'une banque de leur ADN génomique à des cellules murines L929. L'analyse du transcriptome de clones sensibles à l'infection par l'EV-A71 a permis d'identifier le gène de SCARB2 dans le transgène présent dans les cellules



**Figure 21. SCARB2 : protéine membranaire des lysosomes et récepteur d'attachement et d'internalisation de l'EV-A71**

**(A) Structure du récepteur SCARB2.** Il s'agit d'une protéine membranaire de 478 acides aminés qui possède deux domaines transmembranaires, caractérisés par des queues cytoplasmiques en parties N-terminal et C-terminal. Les numéros représentent les exons. La capside de l'EV-A71 peut se lier à l'exon 4 de SCARB2 (d'après Yamayoshi *et al.*, 2012).

**(B) Mécanisme d'infection cellulaire des cellules RD par un EV-A71 via SCARB2.** Physiologiquement, SCARB2 participe aux transports membranaires et à la réorganisation des compartiments lysosomes/endosomes. Lors d'une infection par un EV-A71, SCARB2 peut lier la capside virale et enclencher un mécanisme d'endocytose dépendante de la Clathrine. Le virus est décapsidé et l'ARN viral libéré au sein du cytoplasme cellulaire (d'après Yamayoshi *et al.*, 2012).

**(C) Schéma du mécanisme d'entrée de l'EV-A71 via le récepteur SCARB2.** La capside de l'EV-A71 s'attache à SCARB2 au niveau l'axe d'ordre 5 de la capside virale (A,B). En condition de ph acide (endosome tardif), un changement de conformation de la capside a lieu afin d'ouvrir le canal lipidique (C) et libérer le « pocket factor » représenté par les vers (D,E). L'absence de ce dernier permet la décapsidation et le relargage de l'ARN viral au sein du cytoplasme cellulaire (F ; d'après Dang *et al.*, 2014).

L929 (Yamayoshi *et al*, 2009). La protéine SCARB2 est une protéine intégrée à la membrane lysosomale (famille des CD36) ; elle possède deux domaines transmembranaires [Figure 21a] et c'est l'une des protéines les plus abondantes dans la membrane des lysosomes et des endosomes. SCARB2 intervient dans la réorganisation des compartiments lysosomes-endosomes et agit comme un récepteur pour le transport de la  $\beta$ -glycocerebrosidase ( $\beta$ -GC) du réticulum endoplasmique vers les lysosomes [Figure 21b]. L'expression ubiquitaire de cette protéine dans les tissus humains (y compris les neurones) suggère qu'elle peut être impliquée comme récepteur dans les infections systémiques par l'EV-A71. L'utilisation de protéines chimères entre des formes humaines et murines de SCARB2 a montré que le site de liaison de l'EV-A71 est localisé dans le domaine extracellulaire de SCARB2 (exon 4), entre les acides aminés 142 et 204 [Figure 21a] (Yamayoshi et Koike, 2011). Des acides aminés de la boucle EF de la protéine de capsid VP1 – mur du canyon – sont importants pour la liaison SCARB2/EV-A71 (Chen *et al*, 2012). Le modèle cellulaire L-SCARB2 obtenu par transfection du gène humain de la protéine SCARB2 dans la lignée L929 montre qu'une majorité de souches d'EV-A71 utilisent SCARB2 comme récepteur cellulaire (Yamayoshi *et al*, 2009). L'infection des cellules RD par les sérotypes CV-A7, CV-A14 et CV-A16 est également médiée par cette protéine (Yamayoshi *et al*, 2012).

Contrairement au récepteur PSGL-1, SCARB2 permettrait à la fois l'attachement de la particule virale, son internalisation et sa décapsidation (Yamayoshi *et al*, 2013). L'entrée des virions par endocytose mobiliserait la voie des clathrines. Le mécanisme moléculaire de la décapsidation dépendant de SCARB2 a fait l'objet d'une étude structurale approfondie (Dang *et al*, 2014). Le modèle propose que l'acidification des endosomes soit un élément important permettant un changement de conformation du résidu  $\alpha 5$  de SCARB2. Ceci entraîne l'ouverture d'un tunnel lipidique qui peut faire sortir le lipide appelé « pocket factor » naturellement présent dans les capsides d'entérovirus. La capsid virale change alors de conformation et une ouverture se forme au niveau de la partie N-ter de VP1. Le génome viral est finalement libéré hors de la capsid dans le cytoplasme cellulaire [Figure 21c].

### **3) L'annexine II**

L'annexine II est présente dans plusieurs compartiments (cytosol, membrane). Elle est multifonctionnelle : elle joue un rôle dans l'angiogenèse, la migration cellulaire et la synthèse de plasmine. Elle est notamment exprimée par les cellules endothéliales, les monocytes, les macrophages, les cellules dendritiques et les cellules épithéliales. Présente à la surface des cellules RD, l'annexine II peut se lier à la protéine de capsid VP1 de l'EV-A71 facilitant

l'infection des cellules RD (Yang *et al*, 2011). Cette protéine se lie à des acides aminés compris entre les positions 40 à 100 de la protéine VP1 de l'EV-A71. Cette interaction semble favoriser l'infection virale dans des conditions défavorables pour le virus telles qu'une faible dose infectieuse. Le sérotype CV-A16 ne se lie pas à l'annexine II (Yang *et al*, 2011).

#### **4) L'acide sialique**

Les acides sialiques sont des monosaccharides présents dans les chaînes de glycanes des glycolipides et des glycoprotéines. Ils sont présents en abondance à la surface des cellules épithéliales intestinales. Les acides sialiques de type « N-linked Glycan » et « O-linked Glycan » sont impliqués dans l'infection des cellules épithéliales intestinales DLD-1 par l'EV-A71 (Yang *et al*, 2009).

#### **5) La lectine DC-SIGN « DC-specific intercellular adhesion molecule-3 grabbing nonintegrin »**

Les cellules dendritiques sont des cellules professionnelles présentatrices d'antigènes qui activent les lymphocytes T et induisent la sécrétion de cytokines. L'EV-A71 peut infecter les cellules dendritiques immatures humaines et induire l'expression d'antigènes viraux à la surface de ces cellules infectées (Lin *et al*, 2009). L'entrée de l'EV-A71 dans les cellules dendritiques est en partie médiée par la molécule de surface DC-SIGN. L'infection des cellules dendritiques par l'EV-A71 pourrait ensuite induire la réponse immunitaire adaptative en stimulant la prolifération des cellules T. Bien que l'attachement de l'EV-A71 médié par DC-SIGN ait été confirmé avec des cellules dendritiques dérivées de monocytes (MDCCs), il n'est pas suivi d'une infection virale (Ren *et al*, 2014). La liaison de l'EV-A71 à la molécule DC-SIGN des cellules dendritiques pourrait constituer un mode de dissémination précoce du virus dans l'organisme.

## **IV. Modèles animaux de l'infection à EV-A71**

### **1) Modèle simien**

Les primates de plusieurs espèces sont sensibles à l'infection à EV-A71 : macaque rhésus (*Macaca mulatta*), singe vert (primate simiiforme, *Chlorocebus sabaesus*) et macaque crabier ou macaque cynomolgus (*Macaca fascicularis*). Chez ces primates, l'infection conduit à un syndrome neurologique proche de celui observé chez l'être humain : paralysie flasque



aiguë, troubles locomoteurs et encéphalite du tronc cérébral (Hashimoto *et al*, 1978 ; Koroleva *et al*, 2013). Outre les aspects éthiques de l'utilisation des primates, l'intérêt des modèles simiens d'infection à EV-A71 est limité par le fait que les singes ne développent pas l'œdème pulmonaire qui en est le facteur le plus sévère. Une étude de Zhang *et al* (2011) chez le singe rhésus décrit le mode d'évolution de la virémie dans différents tissus selon le mode d'inoculation.

## 2) Modèle « rongeur »

Les sigmodons, les hamsters dorés et les souris sont les rongeurs qui au stade nouveau-né sont sensibles à l'infection à EV-A71 (Chumakov *et al*, 1979 ; Koroleva *et al*, 2013). Les jeunes sigmodons infectés par des souches d'EV-A71 développent des manifestations neurologiques équivalentes à celles observées chez les singes et chez l'homme ; ils représentent un modèle pour étudier la neuropathologie causée par l'EV-A71 (Koroleva *et al*, 2013). Des lésions pathologiques du SNC (destructions neuronales et gliales) ont été observées suite à l'infection. Concernant les hamsters dorés, des tests d'infections ont été réalisés en 1979, conduisant à des lésions de type poliomyélitique au niveau du SNC, des myosites et des paralysies (Chumakov *et al*, 1979). Mais depuis aucune autre étude n'a repris ce modèle.

Bien que les souris adultes semblent résistantes aux infections par des EV-A71, les souris nouveau-nées y sont sensibles durant une courte période de temps : 2 semaines d'âge au maximum (Wang *et al*, 2014). Une étude indique que les souris de moins de deux semaines ne possèdent pas de récepteurs à l'interféron de type I et II (Khong *et al*, 2012). Des souches d'EV-A71 ont été adaptées à la souris. La souche « MP4 » a été obtenue en effectuant plusieurs passages d'une souche humaine d'EV-A71/C2 chez des souriceaux (Wang *et al*, 2004). Cette souche a révélé des caractéristiques physiopathologiques intéressantes : croissance plus rapide, effet cytotoxique plus important *in vitro* et des atteintes du SNC. Génétiquement, elle a accumulé des mutations protéiques : K218L (protéine VP2) ; A1214P et L1215R (protéine 2C), suggérant des déterminants moléculaires d'adaptation de l'EV-A71 au modèle murin (Wang *et al*, 2004). Mais les études réalisées avec un modèle murin indiquent que les sites principaux d'infection du virus sont le muscle et les tissus adipeux, ce qui en limite l'intérêt physiopathologique. Par ailleurs, les homologues murins des récepteurs humains SCARB2 et PSGL-1 de l'EV-A71 ne jouent pas ce rôle chez la souris. L'équipe japonaise qui a découvert le récepteur SCARB2 a développé des souris transgéniques « hSCARB2-Tg Mice » *via* micro injection dans des embryons murins d'un

**Tableau 4. Principales épidémies causées par l'EV-A71**

Période	Région	Nombre de cas	Nombre de cas sévères	Nombre de décès	Syndromes cliniques	Génotype
1969 - 1972	Californie (USA)	20	20	1	Méningites, encéphalites	A
1972 - 1977	New York (USA)	28	19	0	Méningites, encéphalites, PMB	ND
1972 - 1973	Australie	49	49	0	Méningites	ND
1973	Suède	195		0	Méningites, PMB	ND
1973	Japon	335	84	0	Méningites, PMB	ND
1975	Bulgarie	705	705	44	Méningites, paralysies	ND
1978	Hongrie	1550	1550	45	Méningites (826), encéphalites (724)	ND
1979	France	5	?	0	Atteintes du SNC, fièvre	ND
1986	Australie	144	7	0	Encéphalites (6), paralysies (1)	ND
1986	Hong-Kong	5	5	0	Monoplégie (5)	ND
1997	Sarawak (Malaisie)	6000	?	29	Méningites, paralysies, ataxies cérébrales, œdèmes pulmonaires (29)	B2
1997 - 1998	Singapour	39	?	0	Méningites, PMB	ND
1998	Taiwan	130000	405	78	Encéphalites, œdèmes pulmonaires, PMB	C2
1999	Perth (Australie)	6000	29	0	Méningites, paralysies, ataxies cérébrales, PMB	
2000	Singapour	81	4	4	Méningites, paralysies, PMB	B4
2000 - 2003	Malaisie	277	52	4	Méningites, PMB	B4, B5, C1
2005	Viêt-Nam	173	51	3	Atteintes du SNC (51), PMB (173)	C1, C4
2006-2007	France	28	1	1	Méningites, syndrome respiratoire (1), PMB	C1, C2
2007	Allemagne	13	0	0	Méningites, PMB	C2
2007	Danemark	17	0	0	Méningites, PMB	B5, C1, C2
2007	Shandong (Chine)	1149	?	3	œdèmes pulmonaires, PMB	C4
2008	Chine	488955	?	128	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2009	Chine	666570	?	225	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2010	Chine	1734457	?	905	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2011	Chine	1512064	?	506	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2012	Chine	2071237	?	550	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2012	Cambodge	?	61	61	Atteintes du SNC	C4
2012	Viêt-Nam	148366	?	45	Atteintes du SNC, PMB	ND
2013	Chine	1855078	?	265	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4
2014 (15/07/14)	Chine	1688640	?	220	Méningites, encéphalites, œdèmes pulmonaires, PMB	C4

transgène BAC contenant le gène humain SCARB2 (Fujii *et al*, 2013). Ce modèle est intéressant car il reproduit le neurotropisme observé chez l'homme (ataxie, paralysies et décès). Il a également l'avantage de permettre l'infection par des souches de CV-A16 et de CV-A14.

Comme le modèle simien, la souris ne semble pas constituer un bon modèle pour ce processus, car le profil des cytokines produites est différent de celui observé chez l'Homme (Wang et Yu, 2014).

## **V. Épidémiologie des infections à EV-A**

Les cas de maladie pied-main-bouche surviennent sur un mode épidémique saisonnier avec un premier pic à la fin du printemps ou au début de l'été et un second pic d'intensité moindre au début de l'automne (Solomon *et al*, 2010). La diminution de l'activité épidémique durant l'été pourrait être une conséquence de la fermeture des écoles liée aux vacances scolaires, limitant localement la transmission interindividuelle des souches d'EV-A (Wang *et al*, 2011). La température et d'autres facteurs environnementaux (hygrométrie) pourraient aussi influencer la survie dans l'environnement et la transmission des souches d'EV-A (Wang *et al*, 2011).

### **1) Épidémiologie des infections à EV-A71**

#### **A. Souche prototype**

L'EV-A71 a été isolé pour la première fois pendant une épidémie en Californie en 1969, au cours de laquelle 20 patients (2 mois–46 ans) ont présenté des troubles du SNC : encéphalite, méningite aseptique ou méningoencéphalite. L'analyse des souches par neutralisation croisée a indiqué qu'il s'agissait d'un sérotype distinct de ceux mis en évidence auparavant. La souche « BrCr » a été choisie comme souche représentative de ce nouvel entérovirus parmi celles qui avaient été obtenues pour 20 prélèvements, car sa capacité de multiplication dans des cellules diploïdes de rein humain (HFDK) limitait les risques de contamination par des virus simiens. La souche BrCr a été isolée à partir des selles d'un nourrisson de 2 mois qui avait présenté une méningite aseptique (Schmidt *et al*, 1974).

#### **B. Premières vagues épidémiques, années 1970–1990 [Tableau 4]**

L'EV-A71 a été identifié à New York entre 1972 et 1977, 28 cas dont 19 infections du SNC – encéphalites, paralysies (Deibel *et al*, 1975). Des manifestations neurologiques

causées par l'EV-A71 ont été rapportées à la même période (1972–1973) en Australie chez 49 patients (Kennett *et al*, 1974). En 1973, au sud de la Suède, 195 souches d'EV-A71 ont été isolées chez des patients présentant une méningite aseptique ou une maladie PMB (Blomberg *et al*, 1974). Au Japon, 335 cas d'infections à EV-A71 ont été rapportés pendant une épidémie de maladie PMB ; des atteintes du SNC ont été rapportées dans 25% des cas (Ishimaru *et al*, 1980). En Bulgarie (1975), une épidémie de 705 cas d'infections du SNC a été causée par l'EV-A71 (21% de paralysie flasque de type poliomyélite ; 44 décès ; Chumakov *et al*, 1978 ; Melnick, 1984). Aucun cas de syndrome PMB n'a été rapporté pendant l'épidémie. Une importante épidémie s'est produite trois ans plus tard en Hongrie, 826 cas de méningites aseptiques, 724 cas d'encéphalites et 45 décès (Nagy *et al*, 1982). En 1979, 5 cas d'infections à EV-A71 ont été rapportés parmi des enfants de 5 à 9 ans en France (WHO, 1979). En 1986, au cours d'une épidémie de 144 infections à EV-A71 en Australie, six encéphalites et une paralysie flasque de type poliomyélitique ont été décrites (Gilbert *et al*, 1988). Cinq cas de monoplégie (paralysie localisée) associés à l'EV-A71 ont été rapportés la même année à Hong-Kong (Samuda *et al*, 1987).

### **C. Epidémies dans la région Asie Pacifique depuis 1997**

Dans la période récente, les épidémies de grande ampleur causées par l'EV-A71 ont été rapportées dans la région Asie Pacifique. En 1997, au cours de l'épidémie dans la région de Sarawak (Malaisie), 2 628 cas de syndrome PMB et d'herpangine ont été rapportés (Chan *et al*, 2000). Chez 29 enfants décédés, l'infection à EV-A71 était associée une encéphalite du tronc cérébral et un œdème pulmonaire, une manifestation jamais décrite auparavant pour ce virus. Au cours de l'épidémie de Taiwan en 1998 (130 000 cas de syndrome PMB, 405 cas d'infections du SNC), 78 décès ont été rapportés, principalement liés à un œdème pulmonaire (Ho *et al*, 1999). En 1999, le virus a été lié à 6000 cas de syndrome PMB dans la ville de Perth (Australie) ; 29 cas d'atteintes neurologiques (méningite, ataxie cérébrale et paralysie) ont été notés (McMinn *et al*, 2001). De nombreuses épidémies d'infections à EV-A71 ont été décrites au cours des années suivantes, dans les pays de la région Asie-Pacifique : Singapour, Japon et Corée du Sud en 2000 (Singh *et al*, 2002 ; Fujimoto *et al*, 2002 ; Jee *et al*, 2003), Malaisie en 2000 et 2003 (Podin *et al*, 2006), Viêt-Nam en 2005 (Tu *et al*, 2007) et Brunei en 2006 (AbuBakar *et al*, 2009).

Depuis 2007, les épidémies de maladie PMB les plus importantes sont décrites en Chine. Entre janvier et mai 2008, 61 459 cas de syndrome PMB et 36 décès ont été rapportés. Ceci a conduit à la création d'un système de surveillance du syndrome PMB à l'échelle

nationale en mai 2008. Le syndrome PMB est devenu une maladie de catégorie C dont les cas doivent faire l'objet d'une déclaration (WHO, 2008). Ceci permet la détection d'épidémies multiples dans de nombreuses régions du Nord et du Centre de la Chine, recensant au total 488 955 cas et 126 décès (Yang *et al*, 2009). Les épidémies régionales se sont poursuivies dans le pays en 2009 : 1 155 525 cas dont 13 810 infections du SNC et 353 décès (Yang *et al*, 2011). Elles ont persisté en 2010, 2011 et 2012 (environ 1,5 millions de cas par an). Un bilan fait état de 7 200 092 cas de syndrome PMB déclarés entre 2007 et 2012 (tous sérotypes confondus) dont 82 846 cas d'infections du SNC et 2 457 décès causés principalement par une défaillance cardio-pulmonaire et tous liés une infection à EV-A71 (Xing *et al*, 2014). Les enfants entre 12 et 23 mois ont été les plus affectés par le syndrome PMB pendant les épidémies annuelles dont le pic d'incidence principal est dans la période printemps/été (Xing *et al*, 2014). En 2013, 1 855 078 cas de syndrome PMB ont été déclarés en Chine dont 265 décès. Le bilan établi au 15 juillet 2014 rapporte 1 688 640 cas de syndrome PMB et 220 décès, montrant la persistance du virus dans le pays ([http://www.wpro.who.int/emerging\\_diseases/HFMD/en](http://www.wpro.who.int/emerging_diseases/HFMD/en)).

En 2012, une épidémie au Cambodge a provoqué 61 décès dont 46 morts dans les 24 heures qui suivaient l'admission en milieu médicalisé et les autres dans les 3 jours suivants. Les enfants ont développé des atteintes neurologiques (Biswas, 2012). L'institut Pasteur au Cambodge a identifié l'EV-A71 comme l'agent responsable de cette épidémie. Un réseau de surveillance de syndromes neuro-respiratoire a été mis en place dans le pays. Une épidémie a eu lieu au Viêt-Nam à la même période, 148 366 cas de syndromes PMB dont 45 décès (Khanh *et al*, 2012). En Australie (2013), une épidémie d'infections à EV-A71 a été décrite chez 37 enfants à Sidney (Horsley *et al*, 2014).

Dans le reste du monde, l'EV-A71 circule à bas bruit en Amérique, en Afrique et en Europe, causant des cas sporadiques (Solomon *et al*, 2010). Entre 2003 et 2005, une épidémie d'infections par des EV-A71 causant 16 cas sévères a eu lieu à Denver, aux USA (Pérez-Vélez *et al*, 2007). Au Brésil, une étude récente montre la circulation du virus dans la population (Luchs *et al*, 2010). En Afrique, deux petites épidémies ont été rapportées au Kenya à Nairobi en 1999 et 2000, chez des malades infectées par le VIH (Chakraborty *et al*, 2004). Une étude récente montre la présence de variants d'EV-A71 en Afrique (Nigéria, Cameroun, République Centrafricaine) et Madagascar (Bessaud *et al*, 2014).

## **D. Epidémies et circulation de l'EV-A71 en Europe**

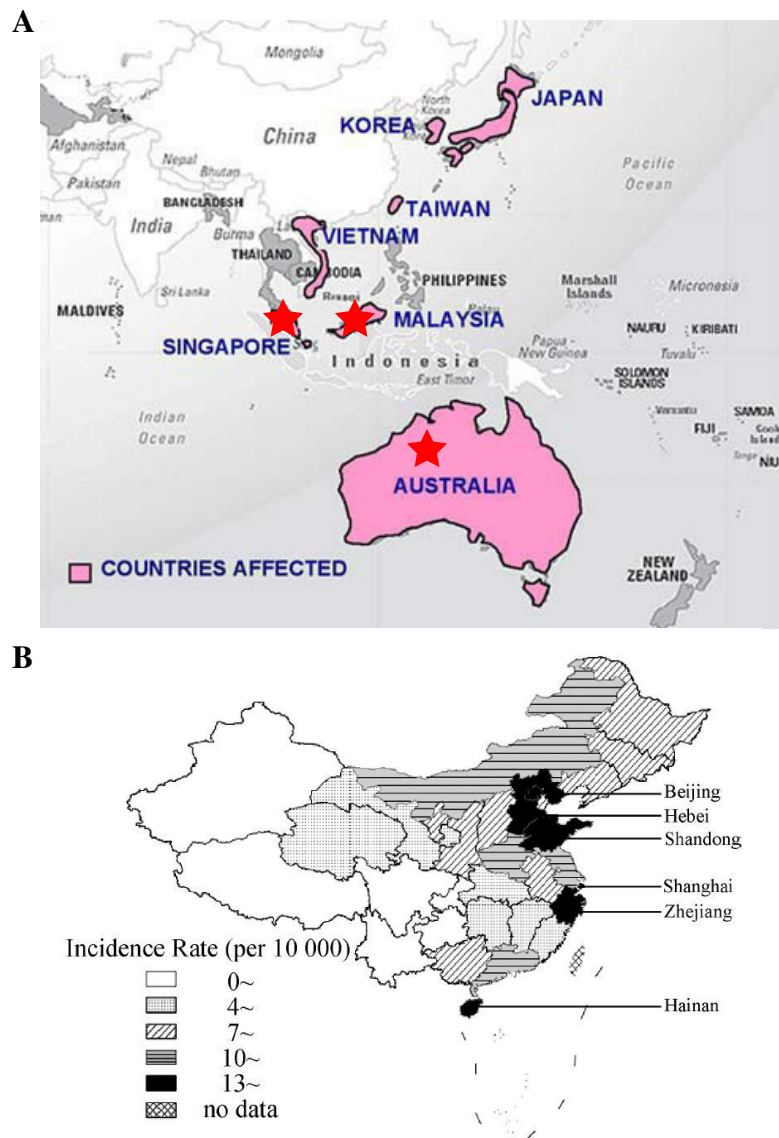
Aucune épidémie de grande ampleur n'a été rapportée en Europe depuis les deux épidémies survenues pendant les années 1970. Plusieurs études ont décrit des cas sporadiques en 1998 en Allemagne (Kehle *et al*, 2003), à Chypre en 2000 et 2002 (Richter *et al*, 2006), en Norvège en 2002 et 2003 (Witsø *et al*, 2007), en Autriche entre 2001 et 2004 (Huemer *et al*, 2008) et en Hongrie entre 2000 et 2005 chez des patients présentant une encéphalite (Kapusinszky *et al*, 2010). En Grande Bretagne, 32 cas d'infection à EV-A71 ont été détectés chez des patients en 1998 et 2006 ; 13 cas ont été associés à une manifestation neurologique dont un décès par encéphalite (Bible *et al*, 2008). Entre 2000 et 2007, 26 souches d'EV-A71 ont été isolées en Allemagne dans le cadre de la surveillance des entérovirus (Diedrich *et al*, 2009).

L'étude d'épidémiologie moléculaire réalisée sur 199 souches d'EV-A71 isolées aux Pays-Bas entre 1963 et 2008 indique la survenue de deux épidémies en 1986 et en 2007 (van der Sanden *et al*, 2009). Elle montre également la circulation de l'EV-A71 dès 1963 en Europe et de ce fait, une circulation antérieure à l'isolement de la souche prototype BrCr en 1969. Le Danemark a également connu une petite épidémie causée par l'EV-A71 en 2007 (Badran *et al*, 2011). Le décès d'un nourrisson de 17 mois causé par un œdème pulmonaire et une rhombencéphale a aussi été rapporté en 2007 en France en lien avec une infection à EV-A71 (Vallet *et al*, 2009). Une étude suggère la circulation active des EV-A en Russie, notamment dans les régions limitrophes de la Chine (Akhmadishina *et al*, 2014).

L'étude réalisée par notre équipe en 2011 montre la circulation de souches d'EV-A71 en France, en rapportant les cas de 55 patients hospitalisés entre 1994 et 2009 (Schuffenecker *et al*, 2011). Les complications neurologiques observées chez 16 patients étaient les suivantes : méningites (n=11), méningo-encéphalites (n=3), œdème pulmonaire (n=1), myélite (n=1) ; deux patients sont décédés. Cette étude a mis en évidence une plus grande fréquence des infections en 2006 et 2007 en France comparativement aux années antérieures.

### **2) Epidémiologie des infections causées par les autres EV-A**

L'épidémiologie des sérotypes CV-A2, CV-A3, CV-A4, CV-A5, CV-A6, CV-A8, CV-A10 et CV-A14 est moins bien connue car ils sont rarement à l'origine de complications sévères (Chang *et al*, 1999). Mais la surveillance renforcée de la maladie pied-main-bouche en Asie montre la circulation épidémique des autres sérotypes dans plusieurs études récentes. Par exemple, des épidémies de maladie pied-main-bouche causées par le CV-A6 ont été rapportées dans différents pays asiatiques : à Taiwan en 2010 (Wei *et al*, 2011), au Japon en



**Figure 22. Réseaux sentinelles de surveillance de la région Asie-pacifique**

**(A) Couverture du réseau de surveillance APNET.**

(<http://sydney.edu.au/medicine/apnet/about/surveillance/index.php>).

**(B) Etendu du réseau de surveillance sentinelle des syndromes PMB en Chine. La carte représente l'incidence des cas de syndromes PMB en Chine en 2009 (d'après Zhu *et al.*, 2011).**

2011 (Kobayashi *et al*, 2013), en Thaïlande en 2012 (Puenpa *et al*, 2013), en Chine en 2013 (Han *et al*, 2014) et en Nouvelle Zélande depuis 2012 (Hayman *et al*, 2014). Sur le continent Américain, la circulation du CV-A6 a été rapportée dans de nombreux états aux USA ainsi qu'à Cuba (CDC, 2012 ; Fonseca *et al*, 2014). Les souches de CV-A6 sont détectées chez des enfants mais aussi des adultes et des signes cliniques atypiques (notamment des lésions cutanées sur l'intégralité du corps) sont rapportés (Mathes *et al*, 2013). Concernant l'Europe, des épidémies de maladie pied-main-bouche associées aux sérotypes CV-A6 et CV-A10 ont été décrites en Espagne, en France et en Finlande entre 2008 et 2012 et en Ecosse en 2014 (Blomqvist *et al*, 2010 ; Mirand *et al*, 2012 ; Montes *et al*, 2013 ; Sinclair *et al*, 2014).

Concernant le sérotype CV-A16, la souche prototype « G10 » a été isolée en 1949 en Afrique du Sud. Les souches de ce sérotype sont souvent détectées lors des épidémies à EV-A71. La surveillance des entérovirus au Japon entre 1985 et 2002 rapporte des vagues d'épidémies d'infections à CV-A16 tous les 3 ou 4 ans (Hosoya *et al*, 2007 ; Mizuta *et al*, 2013). Le virus est détecté depuis 1997 à Taiwan, en Malaisie et au Viêt-Nam (Chang *et al*, 1999 ; Perera *et al*, 2007). La fréquence élevée de la détection de ce sérotype en Chine suggère une circulation plus intense, peut-être endémique à l'échelle globale du pays. Depuis 2007, le sérotype CV-A16 est responsable de la majorité des cas lors des épidémies annuelles de maladie pied-main-bouche en Chine (Zhang *et al*, 2010 ; Xing *et al*, 2014). En Europe, des souches de CV-A16 sont isolées de manière sporadique dans différents pays tels la Hongrie, l'Allemagne et la France (Kapusinszky *et al*, 2010 ; Rabenau *et al*, 2010 ; Mirand *et al*, 2012).

### **3) Surveillance de la maladie pied-main-bouche**

#### **A. Epidémiologie-surveillance dans les pays de la région Asie-Pacifique**

Le maintien de l'activité épidémique de l'EV-A71 depuis sa réémergence dans la région Asie Pacifique il y a un peu moins de 20 ans représente un problème de santé publique dans les pays concernés. Les épidémies à EV-A71 en Asie évoquent à certains égards celles de poliomyélite il y a un siècle et la pathogénie des infections (neurotropisme, atteintes cardio-pulmonaires) font de ce sérotype l'entérovirus le plus menaçant après le poliovirus.

Ces caractéristiques ont conduit à la création de réseaux de surveillance sentinelle dans de plusieurs pays de la région Asie-Pacifique (Taiwan en 1998, Corée du Sud et Viêt-Nam en 2005, Chine en 2007). Une collaboration entre cliniciens, épidémiologistes et virologistes de Singapour, Sarawak (Malaisie) et Sidney (Australie) a donné naissance à APNET (Asia-Pacific Enterovirus Surveillance Network) en 2003 [Figure 22A]. Ce réseau a pour but de



centraliser et traiter les informations épidémiologiques de tous les pays de la région Asie-Pacifique concernant les entérovirus afin de détecter les épidémies. Il informe les services de santé de tous les pays concernés en cas de recrudescence de l'activité virale. Les recherches menées dans les laboratoires faisant partie de ce réseau concernent le développement d'outils rapides de diagnostic des EV-A et également l'étude de la pathogénèse des EV-A. Il assure également la formation et le support de nombreux services de santé impliqués dans la surveillance régionale de l'EV-A71. Il participe également au développement d'un vaccin contre l'EV-A71. En raison des épidémies annuelles de grande ampleur depuis 2007, la Chine a également mis en place des réseaux sentinelles de surveillance du syndrome pied-main-bouche dans les plus grandes villes du pays [Figure 22B].

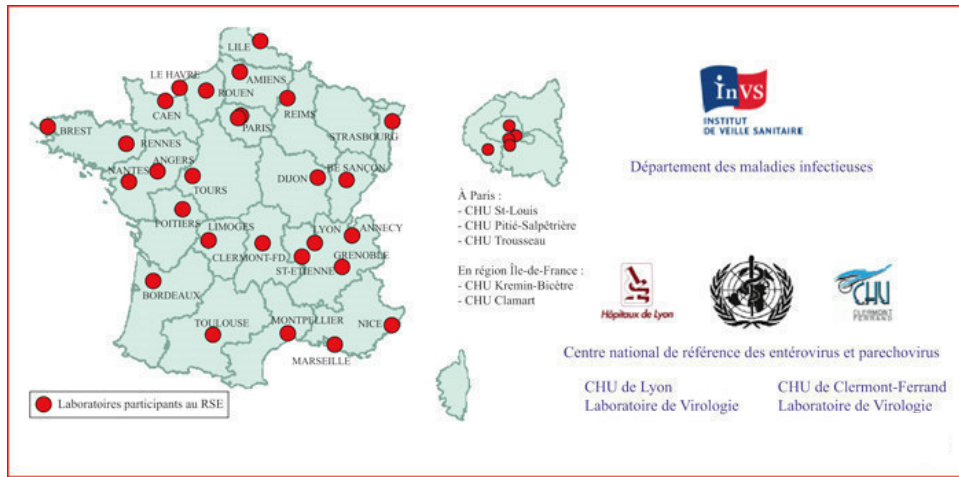
### **B. Epidémiologie-surveillance des infections à entérovirus aux USA et en Europe**

Ailleurs dans le monde, la maladie pied-main-bouche ne fait pas l'objet d'une surveillance spécifique. Plus généralement, la surveillance des infections à entérovirus (excepté la poliomyélite qui fait l'objet d'un programme spécifique coordonné par l'OMS) est intégrée dans les systèmes nationaux de veille sanitaire des maladies infectieuses propres à chaque pays. Aux USA, le programme de surveillance des entérovirus « National Enterovirus Surveillance System » initié en 1961 par le centre de contrôle des maladies infectieuses (CDC) permet de centraliser les cas d'infections diagnostiquées, d'identifier les souches et de mener des investigations plus poussées dans des situations épidémiques. En Europe, le Centre Européen de Prévention et Contrôle des maladies (European Centre for Disease Prevention and Control, ECDC), une Agence indépendante de l'Union européenne, a pour but de renforcer la lutte contre les maladies infectieuses en Europe. Les infections à entérovirus non poliomyélitiques ne font pas partie des virus à déclaration obligatoire. La surveillance des entérovirus n'est pas réalisée de façon homogène dans les tous pays européens. Toutefois, elle est effectuée par des laboratoires hospitaliers spécialisés et, généralement, le laboratoire national de référence affecté à la surveillance de la poliomyélite prend aussi en charge les infections causées par les autres entérovirus. Un réseau de surveillance mis en place en 1996 sous l'égide du RIVM aux Pays-Bas, a permis de mettre en évidence par exemple une épidémie d'infections à EV-A71 en 2007 (van der Sanden *et al*, 2009).

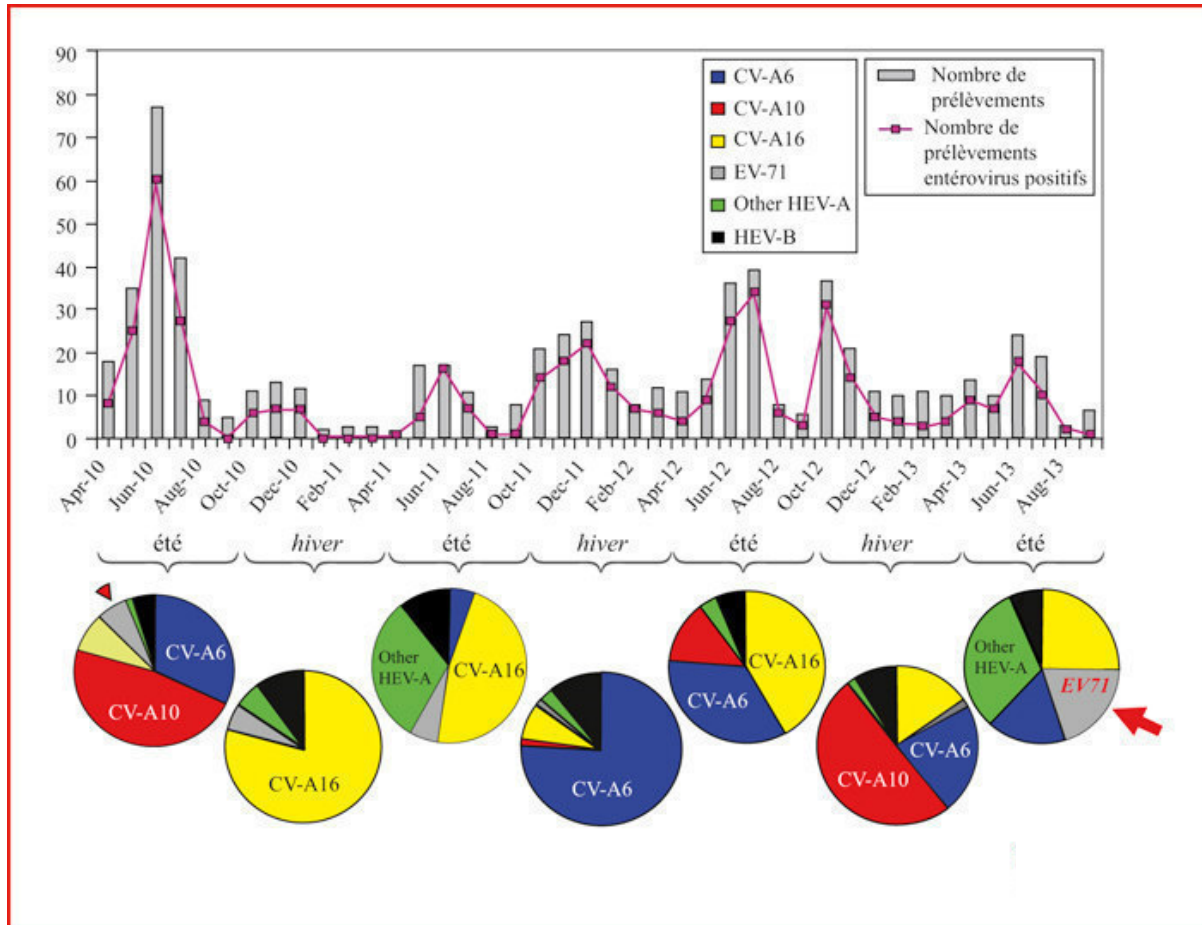
### **C. Epidémiologie-surveillance des infections à entérovirus en France**

En France, la surveillance des infections à entérovirus est réalisée par le Centre National de Référence (CNR) des Entérovirus et Parechovirus qui dépend de l'Institut de

A



B



**Figure 23. Surveillance des entérovirus en France**

(A) Carte des laboratoires participants au RSE en 2013 (d'après Peigue-Lafeuille *et al.*, 2014).

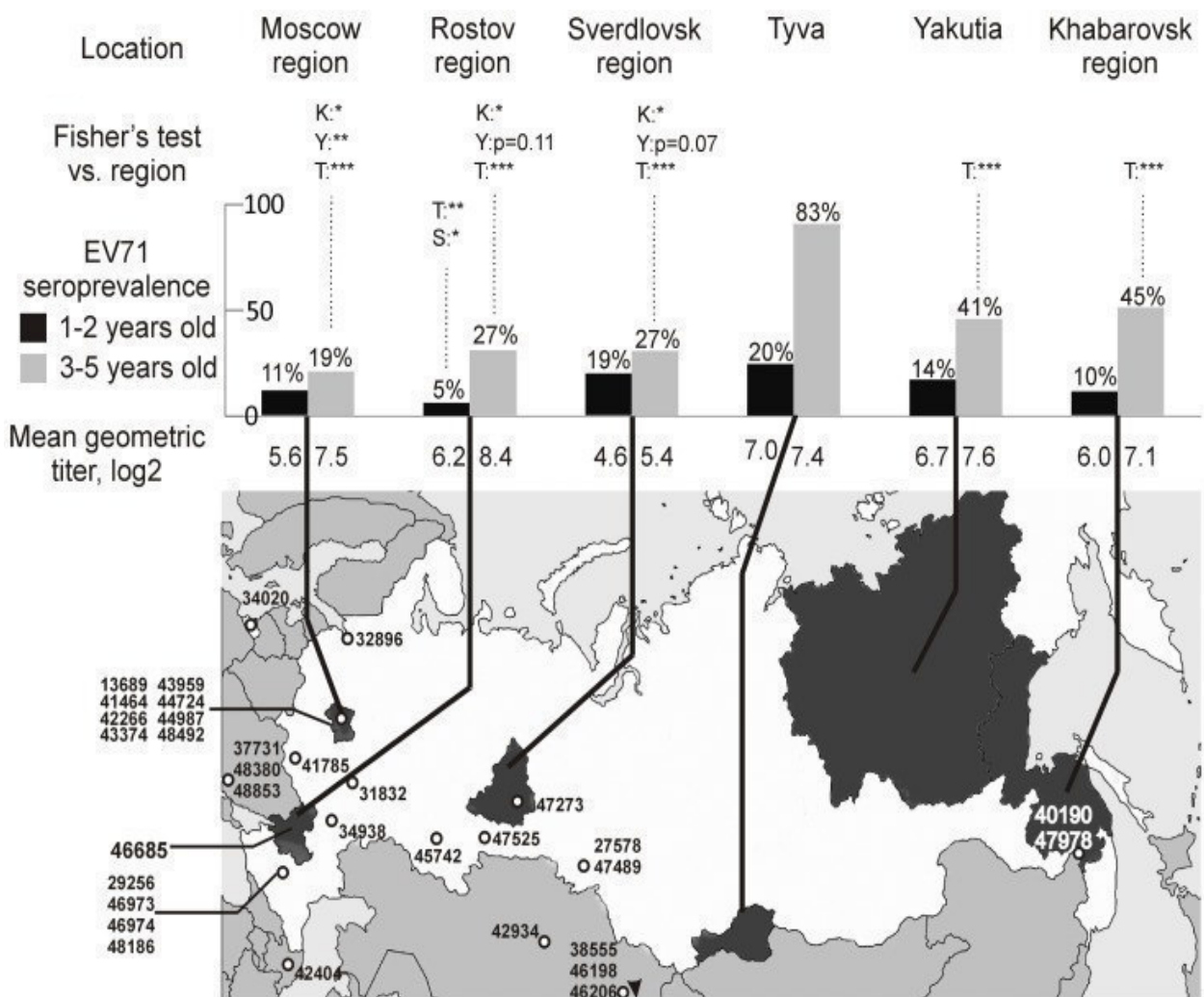
(B) Répartition annuelle du nombre de prélèvements positifs à EV et distribution des sérotypes dans le cadre de la surveillance des syndromes PMB (Clermont-Ferrand, 2010-2012 / 2013).

Veille Sanitaire (InVS). Les données cliniques et épidémiologiques sont obtenues par le système de surveillance et d'alerte mis en place en 1996. Le Réseau de Surveillance des EV (RSE) est actuellement composé de 40 laboratoires hospitaliers qui assurent le diagnostic virologique. Le CNR Entérovirus/Parechovirus centralise les données et assurent un service d'identification virale, de diagnostic le cas échéant et les alertes sanitaires lors de la survenue d'épidémies ou de formes cliniques inhabituelles [Figure 23A].

En 2010, notre équipe a initié la surveillance spécifique des cas de syndrome pied-main-bouche pour l'agglomération de Clermont-Ferrand. Un réseau sentinelle incluant 17 pédiatres du secteur privé et le service des urgences pédiatriques du CHU de Clermont-Ferrand a été constitué pour une étude de trois ans. Le but du réseau est d'effectuer le diagnostic virologique prospectif des cas de maladie PMB et d'herpangine et l'identification des souches dans les prélèvements de gorge ou de bouche réalisés chez les enfants [Figure 23B]. En 2014, ce réseau a été étendu à l'échelle nationale dans le cadre d'un protocole de recherche clinique auquel participent 42 pédiatres libéraux répartis dans la plupart des régions (Corse comprise). Cette étude permettra de comparer les caractéristiques cliniques associées aux différents types d'entérovirus rencontrés à l'échelle nationale. Le nombre potentiel d'enfants inclus sera de 3120 en un an.

## **VI. Séroprévalence de l'infection à EV-A71**

Les études de séroprévalence informent sur le statut immunitaire de la population générale et aussi sur la circulation silencieuse du virus. La première étude à s'être intéressée à la prévalence de l'infection à EV-A71 dans un pays de la région Asie-Pacifique concerne la population de Taiwan avant et après l'épidémie de 1998 (Chang *et al*, 2002). Elle a été effectuée avec une collection de 539 sérums prélevés en 1997 et de 4619 sérums obtenus en 1999. Après l'épidémie (prélèvements effectués entre janvier et juillet 1999), la séroprévalence de l'EV-A71 chez les enfants de moins de moins de 3 ans était comprise entre 8% et 24%. L'étude d'un échantillon de 161 sérums prélevés dans la population thaïlandaise indique une forte prévalence de l'infection à EV-A71 (65,8%) en 2009/2010 (Linsuwanon *et al*, 2014). L'analyse rétrospective de 900 sérums collectés en 2005 dans 6 provinces chinoises indique une grande disparité des taux d'infections à EV-A71 et CV-A16 entre les provinces ; les taux d'infection à EV-A71 étaient inférieurs à 40% chez les enfants de moins de 3 ans avant l'épidémie de 2008 dans trois provinces (Zhu *et al*, 2010).



**Figure 24. Séroprévalence de l'EV-A71 dans six régions de la Fédération de Russie**

La séroprévalence de l'EV-A71 a été étudiée chez des enfants âgés entre 1–2 et 3–5 ans au sein de différentes régions de Russie. Le numéro des souches d'EV-A71 isolées dans ces régions sont indiquées sur la carte. Les différences significatives (test de Fisher) de séroprévalence entre les différentes régions sont indiquées par les lettres au dessus du graphique de séroprévalence. \* :  $p < 0.05$ , \*\* :  $p < 0.01$ , \*\*\* :  $p < 0.001$ .  
 (D'après Akhmadishina *et al.*, 2014).

La séroprévalence de l'infection à EV-A71 a été étudiée avec 831 sérums recueillis chez des enfants de 6 régions russes dont certaines limitrophes de la Chine (Akhmadishina *et al*, 2014). Le taux d'infection varie entre 19 % et 83 % selon la localisation géographique : elle est significativement plus importante chez les enfants de 3 à 5 ans vivant les régions de la partie asiatique [Figure 24].

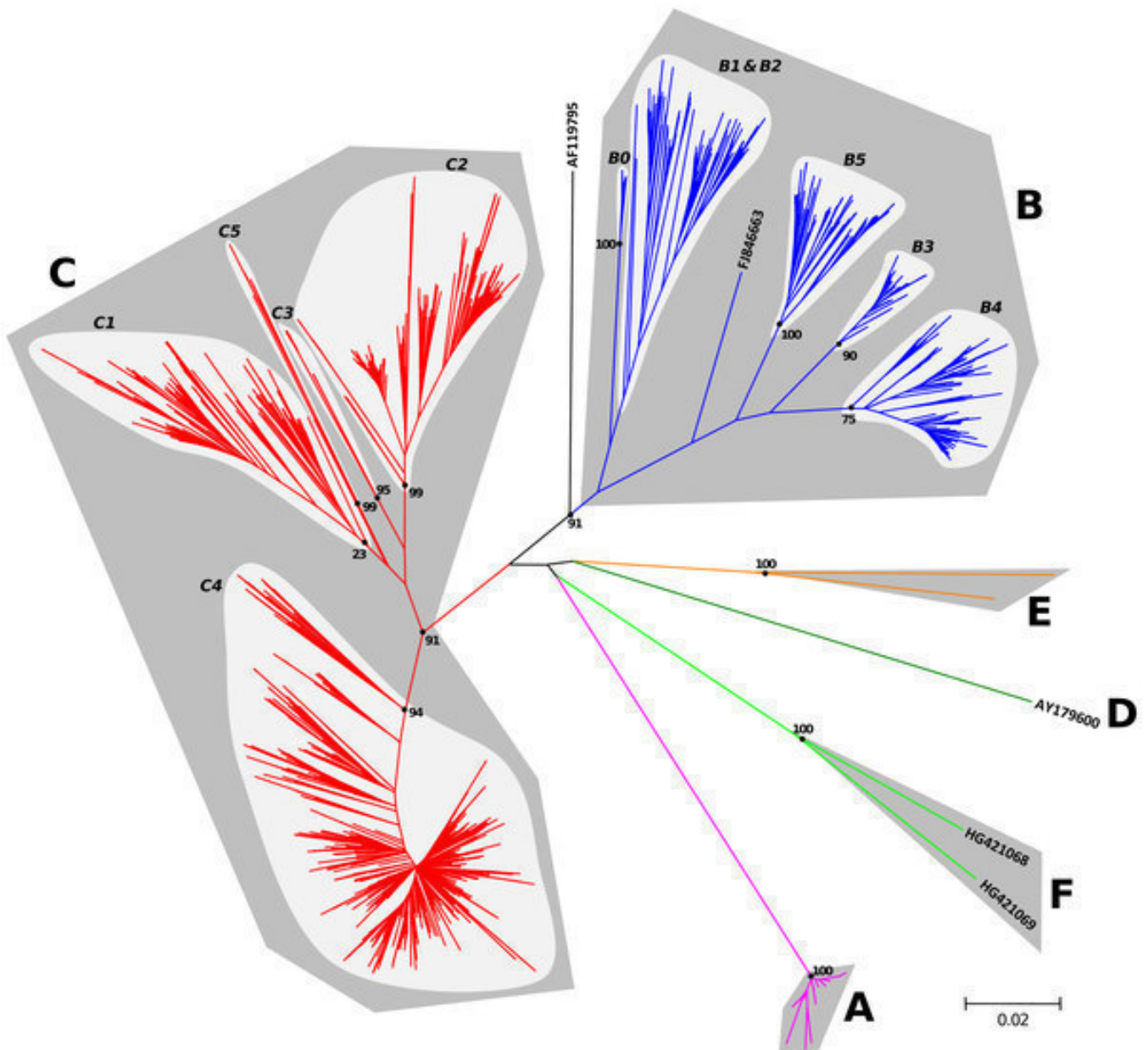
Concernant l'Europe occidentale, deux études réalisées en Allemagne montrent que les enfants d'âge préscolaire sont très majoritairement séronégatifs (Diedrich *et al*, 2009 ; Rabenau *et al*, 2010). Mais il existe des variations importantes entre les deux études : 63.4 % (0–3 ans) dans la première étude et 88 % (0–4 ans) dans la seconde. (Les taux de séroprévalence varient entre 40 % et 50 % chez les 5–9 ans dans les deux études, mais les résultats ne sont pas concordants pour les classes d'âge supérieur : dans la première étude, le taux d'infection chez les individus de 15 ans et plus est >60 % et il oscille entre 40 % et 50 % dans la seconde.

## VII. Epidémiologie moléculaire des entérovirus de l'espèce A

L'épidémiologie moléculaire des entérovirus repose sur l'analyse des séquences structurales du génome car les événements de recombinaison génétique sont très rares dans toute la région génomique P1 (Oberste *et al*, 1999 ; Bouslama *et al*, 2007). La séquence de la protéine de capsid VP1 est la plus utilisée pour génotyper les souches, elle est reliée à plusieurs caractéristiques virales (Bailly *et al*, 2008). Elle porte des déterminants antigéniques majeurs, elle interagit avec un ou plusieurs récepteurs cellulaires et avec les trois autres protéines de capsid. Les séquences de la protéine de capsid VP2 permettent également un génotypage fiable mais sont moins utilisées (Ibrahim *et al*, 2014).

### 1) Epidémiologie moléculaire de l'EV-A71

La séquence nucléotidique du génome complet de la souche prototype BrCr d'EV-A71 a été déterminée en 1995, montrant qu'elle était génétiquement distincte du poliovirus (Brown et Pallansch, 1995). La première étude d'épidémiologie moléculaire concernant l'EV-A71 a porté sur l'analyse de 113 séquences complètes du gène 1D<sup>VP1</sup> de souches isolées entre 1970 et 1998 (Brown *et al*, 1999). L'analyse de cet échantillon a montré l'existence de trois lignées monophylétiques dont les séquences présentaient plus de 15 % de divergence nucléotidique, désignées génogroupes A, B et C. La séquence de la souche prototype BrCr a été assignée au génogroupe A. Les deux autres génogroupes B et C, initialement subdivisés, en sous-



**Figure 25. Diversité génétique de l'entérovirus 71 établie à partir du locus 1D<sup>VP1</sup>**

Le sérotype EV-A71 comprend 6 génogroupes (A à F). Les génogroupes B et C se subdivisent en sous génogroupes (B0 à B5 et C1 à C5). Les distances évolutives ont été déterminées par la méthode de Kimura à 2 paramètres. La taille des branches est proportionnelle à la divergence en nucléotides. Le pourcentage de répliquats de Bootstrap est indiqué pour les nœuds principaux.

(D'après Bessaud *et al.*, 2014).

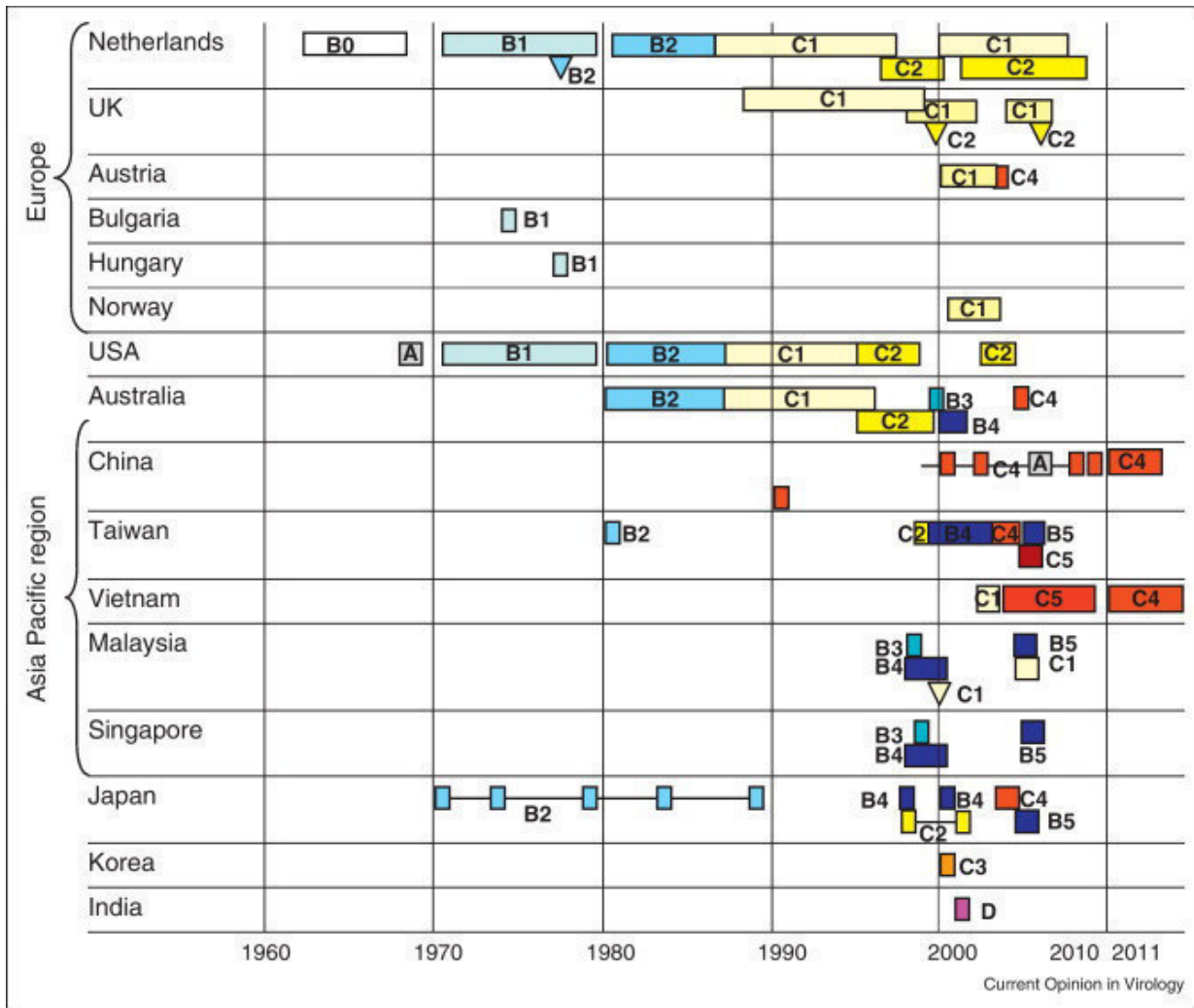
génogroupes (B1, B2) et (C1, C2), ont été étendus aux sous-génogroupes B3, B4, B5, C3, C4 et C5 (Solomon *et al*, 2010). L'analyse rétrospective de souches isolées entre 1963 et 1967 aux Pays-Bas a permis de découvrir le sous-groupe B0, le plus ancien du génogroupe B (van der Sanden *et al*, 2009). La diversité génétique de l'EV-A71 a été étendue à trois autres génogroupes D, E et F après l'identification de souches circulant à Madagascar, en Inde et en Afrique [Figure 25] (Deshpande *et al*, 2003 ; Bessaud *et al*, 2014). Dans une étude très récente, Saxena *et al* (2015) rapportent la caractérisation génétique de souches isolées en Inde et qui indique l'existence d'un génogroupe G supplémentaire.

Bien que la distinction de certains sous génogroupes en utilisant la méthode « neighbour-joining » ne soit pas solide sur un plan phylogénétique, ils peuvent être considérés sur le plan de l'écologie virale (de l'épidémiologie) comme des populations virales plus ou moins transitoires de l'EV-A71. Plusieurs aspects de l'histoire de la circulation du virus peuvent donc être retracés par l'analyse phylogénétique des échantillons de souches de différentes origines géographiques [Figure 26].

Les épidémies rapportées aux USA, en Europe (Hongrie et Bulgarie), au Japon et à Taiwan au cours des années 1960–1970 indiquent la co-circulation des populations B1 et B2 à l'échelle mondiale (Yip *et al*, 2013). Pendant les années 80, le sous-génogroupe B2 est rapporté le plus fréquemment mais n'a pas été relié à des épidémies de grande ampleur. Les souches du génogroupe C1 ont été identifiées en parallèle en Australie, au Japon, aux Pays Bas, sans causer d'épidémies importantes et c'est au cours de la décennie suivante que les populations virales (B3, B4, B5, C2, C3, C4 et C5) ont été identifiées. La co-circulation des populations B3, B4, C1 et C2 a été reliée à deux grandes épidémies de maladie pied-main-bouche (Cardosa *et al*, 2003) : l'épidémie de 1997 en Malaisie fut principalement associée à la population B3 et celle de 1998 à Taiwan à la population C2. La co-circulation des virus B3 et C2 a été notée lors de l'épidémie australienne de 1999 (McMinn *et al*, 2001).

Les sous génogroupes se succèdent lors des épidémies qui surviennent à l'échelle d'un pays, voire simplement d'une région ou localité d'un pays (Huang *et al*, 2009 ; Iwai *et al*, 2009). Certaines populations virales ont une durée de vie très brève et des zones restreintes de circulation. Le sous génogroupe C3 par exemple a été isolé uniquement en Corée du Sud lors d'une épidémie en 2000 (Jee *et al*, 2003) et le variant C5 a été rapporté ponctuellement au Vietnam, en Thaïlande et à Taiwan entre 2005 et 2011 (Huang *et al*, 2009 ; Thao *et al*, 2013).

L'émergence du sous génogroupe C4 est un fait majeur de l'épidémiologie récente de l'EV-A71. Cette population virale est à l'origine d'épidémies annuelles de grande ampleur de



**Figure 26. Distribution géographique des génogroupes et sous-génogroupes de l'entérovirus 71 entre 1960 et 2011**

Les génogroupes et sous génogroupes sont représentés par le code couleur. Les triangles indiquent des isollements sporadiques. (D'après McMin., 2012).



la maladie pied-main-bouche en Chine depuis 2007 et s'est propagée dans toute la région Asie-Pacifique (Zhang *et al*, 2009 ; Yip *et al*, 2013).

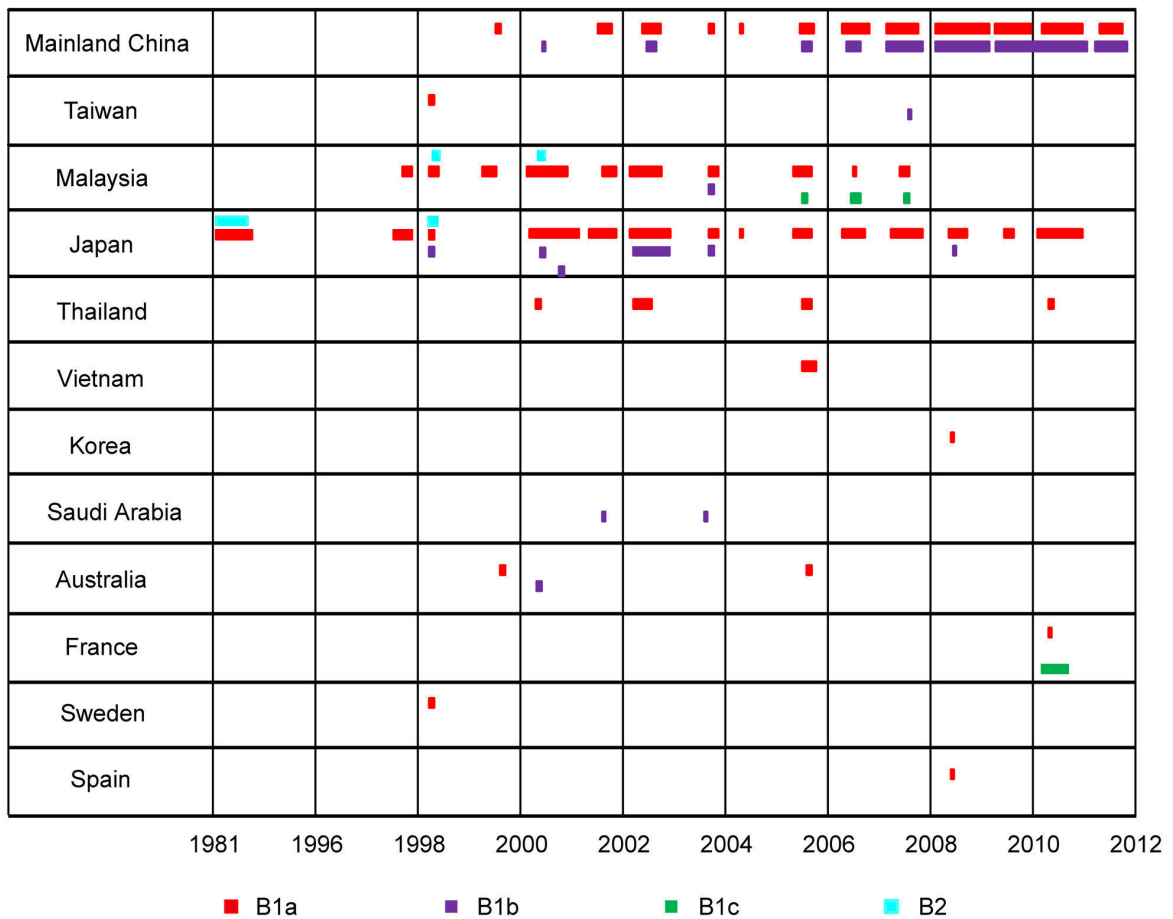
Deux études ont mentionné l'isolement d'une quinzaine de souches du génogroupe A de l'EV-A71 en 2008 et 2009 dans différentes régions Chinoises (Yu *et al*, 2010 ; Zhu *et al*, 2013). Leur origine exacte demeure incertaine. La faible diversité génétique relative de ces souches par rapport à la souche prototype suggère que la circulation récente de cette population virale a été occasionnée à partir d'une souche de laboratoire (Bessaud *et al*, 2014).

L'étude comparée des souches d'EV-A71 anciennes conservées aux Pays-Bas et de souches récentes a montré un « shift » entre le génogroupe B (circulation des populations B0, B1, B2 pendant les années 60–80) et le génogroupe C dont les populations C1 et C2 ont été rapportées pour la première fois à la fin des années 1980 (Van der Sanden *et al*, 2009). Les études épidémiologiques concernant l'Europe indiquent que les populations actuellement en circulation appartiennent aux sous génogroupes C1 et C2 (Van der Sanden *et al*, 2009 ; Mirand *et al*, 2010). Des souches de sous génogroupe C4 ont circulé dans plusieurs pays européens en 2004 : Allemagne, Autriche, France et Hongrie (Huemer *et al*, 2008 ; Schuffenecker *et al*, 2011). Des infections causées par ce variant ont été rapportées en France (quatre cas en 2012) et en Russie (11 cas en 2013), ce qui suggère une circulation en Europe et la nécessité de surveiller ce variant (Akhmadishina *et al*, 2014 ; Schuffenecker *et al*, 2014).

## **2) Epidémiologie moléculaire des coxsackievirus A**

L'investigation clinico-épidémiologique des épidémies de maladie PMB et d'herpangine montre que d'autres entérovirus sont présents avec l'EV-A71, ce qui indique une co-circulation de souches de multiples sérotypes. Les sérotypes de l'espèce EV-A sont très fréquents, ceux de l'espèce B plus épisodiques.

Bien que la majorité des études d'épidémiologie moléculaire coxsackievirus A16 aient porté sur de petits échantillons de souches, l'analyse phylogénétique a permis de définir la structure des populations virales de ce virus, génogroupes et sous génogroupes (Hosoya *et al*, 2007 ; Perera *et al*, 2007 ; Zhang *et al*, 2010 ; Zong *et al*, 2011). L'analyse la plus exhaustive de la diversité génétique globale des souches de CV-A16 a été réalisée sur un échantillon de 607 séquences du locus 1D<sup>VP1</sup> (Chen *et al*, 2013). Le génogroupe A est caractérisé par la souche prototype, le génogroupe B est composé des autres souches en circulation depuis 1981 et il comprend deux sous génogroupes. La population B2 est composée de souches isolées entre 1981 et 2000, et la population B1 contient des souches rapportées depuis 1995 [Figure



**Figure 27. Distribution temporelle et géographique des génogroupes du coxsackievirus A16 entre 1981 et 2012**

Cette distribution a été établie à partir de 628 séquences du locus 1D<sup>VP1</sup> de CV-A16. Les génogroupes sont représentés par le code couleur. La longueur des barres est proportionnelle au nombre d'isolats pour chaque année (d'après Chen *et al.*, 2013).

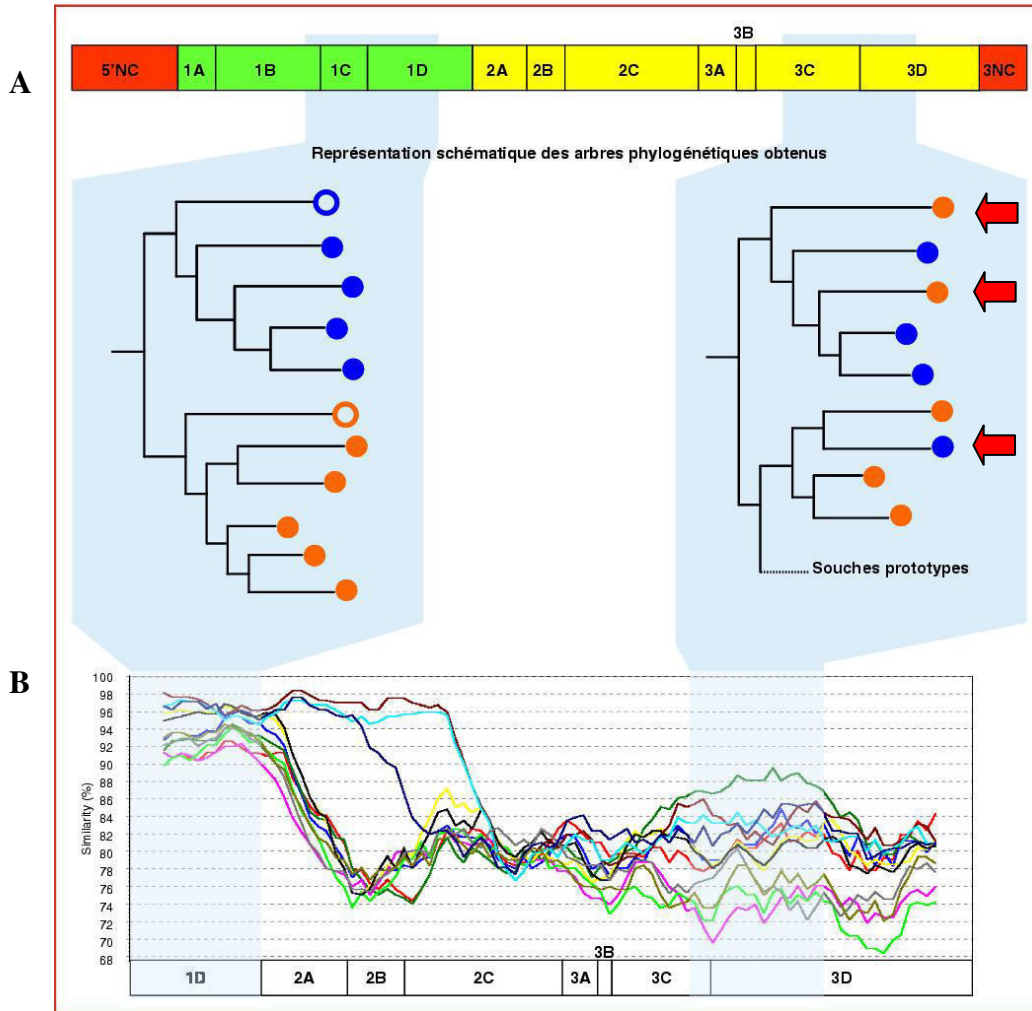
27]. L'épidémiologie du CV-A16 dans des régions géographiques autres que l'Asie est très peu explorée.

D'autres sérotypes de l'espèce A, CV-A6 et CV-A10, ont été à l'origine d'épidémies récentes de syndrome pied-main-bouche chez des enfants et également chez des adultes. Les études de l'épidémiologie moléculaire de ces virus restent limitées aux épidémies locales ou régionales (Mirand *et al*, 2012). L'analyse d'échantillons de séquences partielles (265 nt) de CV-A6 et CV-A10 indique une diversité génétique importante chez ces deux sérotypes et les souches rapportées en Chine sont génétiquement proches de celles qui circulent en Europe (He *et al*, 2013).

### 3) Recombinaison génétique et évolution des entérovirus A

La recombinaison génétique contribue à redistribuer la diversité génétique parmi les souches et elle peut être reliée à l'émergence de nouvelles populations virales circulantes. Ce mécanisme génétique est donc essentiel pour l'épidémiologie et l'histoire évolutive des entérovirus. La recombinaison intratypique est difficile à mettre en évidence dans un échantillon de séquences. La recombinaison intertypique peut être détectée par une approche multilocus en comparant la congruence topologique de deux arbres phylogénétiques estimés pour deux loci distants [Figure 28A] (Lindberg *et al*, 2003). L'analyse de la similitude nucléotidique sur la portion de génome séparant les deux loci permet de confirmer l'existence et de localiser le site de recombinaison [Figure 28B].

La recombinaison génétique parmi les EV-A est un évènement moins fréquent que parmi les EV-B (Simmonds et Welch, 2006). L'analyse des génomes complets des souches prototypes d'EV suggère que certains d'entre eux ont une origine mosaïque par recombinaison intertypique dans les régions codant les protéines non structurales P2 et P3 ; il s'agit des groupes CV-A6/CVA-10/CV-A12 et CV-A4/CV-A14/CV-A16 (Oberste *et al*, 2004). L'analyse multilocus réalisée avec 86 souches d'EV-A71 isolées en Europe (France, Allemagne et Autriche) montre que les topologies des arbres phylogénétiques estimés pour les loci 1D et 3CD ne sont pas congruentes, ce qui suggère que la recombinaison est intervenue dans l'histoire évolutive de plusieurs clades [Figure 29A] (Mirand *et al*, 2010). Une autre étude (McWilliam Leitch *et al*, 2011), réalisée sur 308 souches d'EV-A71 a confirmé ces résultats et propose que des évènements de recombinaison sont liés à l'émergence des sous génogroupes d'EV-A71 [Figure 29B]. Dans ces deux études, aucune analyse de similitude nucléotidique des génomes complets n'a été effectuée afin d'identifier les sites de recombinaison dans les génomes des populations virales en circulation.

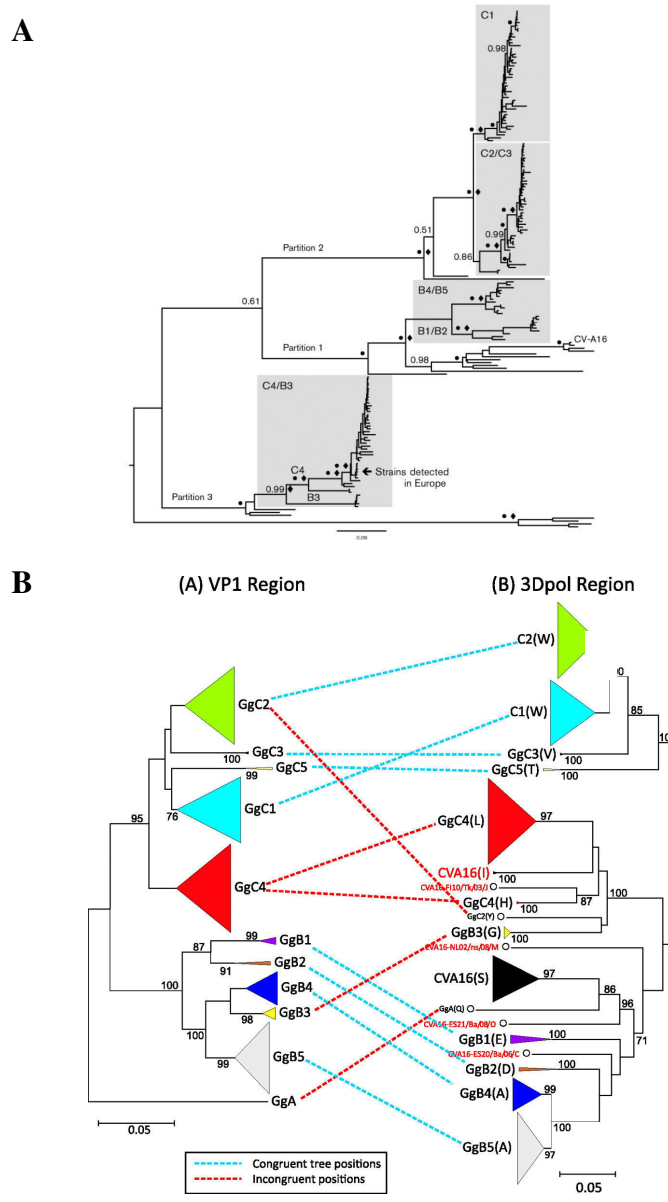


**Figure 28. Détection des souches d'entérovirus d'origine recombinante par une approche multilocus et l'analyse de similitude nucléotidique**

A. Un échantillon de souches d'entérovirus de deux sérotypes (rond bleu, rond rouge) est comparé sur deux locus différents (1D vs 3CD). Les arbres phylogénétiques obtenus par l'analyse de chaque locus présente des différences dans les répartitions de lignées monophylétiques (incongruence de topologie), représentées par les flèches rouges. Ceci suggère la survenue d'évènements de recombinaison à l'origine de ces différences.

B. Une analyse de similitude nucléotidique sur le subgénomme séparant les deux loci permet de confirmer la survenue d'un évènement de recombinaison (diminution brutale de la similitude nucléotidique) et de localiser le site de crossover.

(D'après Bailly et al., 2008).



**Figure 29. Recombinaison de l'EV-A71**

A. Phylogénie établie pour un échantillon de 86 séquences du locus 3CD de souches d'EV-A71. Les symboles indiquent la fiabilité des nœuds (densité de probabilités postérieures > 0,9).  
 (D'après Mirand *et al.*, 2010).

B. Comparaison des topologies phylogénétiques entre les loci 1D<sup>VP1</sup> et 3D<sup>pol</sup> pour un échantillon de 308 souches d'EV-A71. Les incongruences topologiques sont représentées par des pointillées rouges. La fiabilité des nœuds est établie par la méthode du bootstrap.  
 (D'après McWilliam *et al.*, 2011).

Des évènements de recombinaison intertypiques ont été décrits entre l'EV-A71 et le CV-A16 par analyse comparée des génomes (Chan et AbuBakar, 2004 ; Yoke-Fun et AbuBakar, 2006 ; Huang *et al*, 2008). Les critères généralement utilisés pour conclure à une origine recombinante entre deux génomes sont peu discriminants (seuil de similitude nucléotidique inférieur à 90 %). Si l'on considère que le génome d'un entérovirus évolue à un taux d'environ 1 % par an, un tel seuil indique nécessairement une parenté éloignée. Avec un seuil aussi peu stringent, le sous génogroupe C4 serait issu d'une recombinaison entre des souches du génogroupe B et la souche prototype de CV-A16 G10, isolée dans les années 1950 (Yip *et al*, 2010), ce qui semble difficile à concevoir. L'analyse génomique comparée de souches d'EV-A71 isolées entre 1963 et 2010 aux Pays Bas a mis en évidence des sites de recombinaison qui suggèrent que des souches B2 isolées en 1986 aux Pays-Bas présentent des portions de la région 5'NC génomiques communes avec des souches B3 et B4 isolées dans la région Asie- Pacifique (Van der Sanden *et al*, 2011).

Les incongruences\* phylogénétiques et les analyses de similitude nucléotidique entre des génomes de souches isolées en Chine en 2009 montrent de manière fiable la présence de sites de recombinaison entre les génomes de CV-A2 et de CV-A4 dans la séquence génique 2C (Hu *et al*, 2011). Une approche multilocus a été utilisée pour analyser un échantillon de 78 souches d'EV-A isolées en Russie et les pays limitrophes (Lukashev *et al*, 2014). Des indices phylogénétiques de recombinaison récente ou ancienne sont retrouvés pour les souches de tous les sérotypes sauf celles du CV-A16 et de l'EV-A71. Les populations de ces deux sérotypes semblent circuler pendant de longues périodes mais les souches ne portent pas de traces de recombinaison, ce qui suggère une sélection négative des souches d'origine recombinante. Cette étude suggère aussi que les évènements de recombinaison à l'origine de l'émergence des sous génogroupes d'EV-A71 auraient impliqué des souches d'autres sérotypes d'EV-A non décrits (Lukashev *et al*, 2014).

## **VIII. Variabilité génétique et immunité croisée**

La réponse immunitaire humorale constitue un facteur majeur pour limiter la survenue des atteintes sévères pendant les infections à entérovirus. Son rôle protecteur a été bien analysé en comparant l'efficacité relative de la vaccination antipoliomyélitique par les vaccins oraux et inactivés (WHO, 2010). Concernant l'EV-A71, la diversité des populations de ce virus en circulation dans le monde indique que les souches subissent une évolution génétique rapide et

continue qui peut constituer un facteur d'échappement aux réponses immunitaires intra individuelles. La relation entre l'évolution génétique des souches et le degré de séroneutralisation croisée de leur pouvoir infectieux a été explorée parmi les populations d'EV-A71. Ces tests peuvent être réalisés avec des antisérums obtenus chez l'animal de laboratoire ou avec des sérums humains après des infections naturelles.

Des réactions antigéniques croisées ont été mises en évidence entre des souches des génogroupes B et C, ce qui suggère des épitopes communs (Kung *et al*, 2007). Cette observation faite à l'aide d'immuns sérums de lapin a été confirmée en utilisant des sérums de cochons d'Inde (Mizuta *et al*, 2009). Des sérums de lapin anti EV-A71 B2 peuvent neutraliser des souches B0, B1 et B2 mais pas des souches C1 et C2 (Van der Sanden *et al*, 2010). *A contrario*, les anticorps anti C1 peuvent neutraliser les souches des deux génogroupes, ce qui évoque les propriétés des souches « primes », un concept développé par Melnick.

Les sérums de malades ont été utilisés pour comparer les taux de séroneutralisation entre des souches isolées à Taiwan et établir une cartographie des sous génogroupes selon les distances antigéniques déterminées (Huang *et al*, 2009). Les résultats indiquent que les génogroupes B1/B4 et C2/C4 forment des groupes antigéniques différents. Par ailleurs, le sous génogroupe B5 présente des propriétés antigéniques différentes des autres génogroupes d'EV-A71 (Huang *et al*, 2009). Des tests de neutralisation croisée réalisés avec les sérums de 25 enfants taiwanais de moins de 5 ans ayant eu une infection à EV-A71 entre 1998 et 2010 montrent que les anticorps anti C2 et anti C4 neutralisent avec la même sensibilité des souches des génogroupes B et C. Les souches de génogroupe B sont neutralisées plus efficacement que celles du génogroupe C par des immuns sérums anti B4 et anti B5 (Huang *et al*, 2013). La plupart des études soulèvent le manque de standardisation des protocoles utilisés pour tester les propriétés de neutralisation croisée.

## **IX. Prévention des infections et développement d'un vaccin**

Le succès des campagnes de vaccinations pour éradiquer la poliomyélite fait apparaître l'utilisation d'un vaccin contre l'EV-A71 comme un moyen efficace d'enrayer la dissémination de ce virus et de contrôler la survenue d'épidémies chez les très jeunes enfants.

Plusieurs groupes pharmaceutiques en Chine, à Singapour, au Japon, en Malaisie et à Taiwan développent actuellement des vaccins contre l'infection à EV-A71 (Liang *et al*, 2013). Les vaccins candidats sont développés en utilisant différentes lignées cellulaires et des

**Tableau 5. Vaccins inactivés en cours développement contre l'EV-A71**

<b>Organisation</b>	<b>Souche (genogroupe)</b>	<b>Lignée cellulaire</b>	<b>Milieu</b>	<b>Technologie</b>	<b>Adjuvant</b>	<b>Statut d'essais cliniques</b>
<b>Vigoo</b>	FY strain (C4)	Vero	Sans sérum	Système microporteur en bioréacteur	Aluminum Hydroxyde	Phase III achevée
<b>Sinovac</b>	Souche H07 (C4)	Vero	Sans sérum	Incubateur cellulaire	Aluminum Hydroxyde	Phase III achevée
<b>CAMS</b>	Souche M01 (C4)	KMB-17	Avec sérum	Incubateur cellulaire	Aluminum Hydroxyde	Phase III achevée
<b>NHRI</b>	Souche E59 (B4)	Vero	Sans sérum	Bouteilles roulantes	Aluminum Hydroxyde	Phase I achevée
<b>Inviragen</b>	- (B3)	Vero	Sans sérum	Incubateur cellulaire	Aluminum Hydroxyde	Phase I achevée



souches de différents sous génogroupes. Les valences antigéniques qui ont été choisies répondent à des préoccupations épidémiologiques nationales : le sous génogroupe C4 dans les projets en Chine et des souches du génogroupe B dans ceux développés à Singapour et Taiwan. Parmi les modes de préparation vaccinale figurent les vaccins à virus atténué (virus non pathogène), les vaccins de type pseudo-particules virales (capsides vides sans génome dont la structure est similaire à celles des virions infectieux), les vaccins sous-unités à protéines virales recombinantes (protéine VP1) et les vaccins à virus inactivé. Les projets de vaccins basés sur les trois premières approches sont encore en phase de développement sur des modèles animaux.

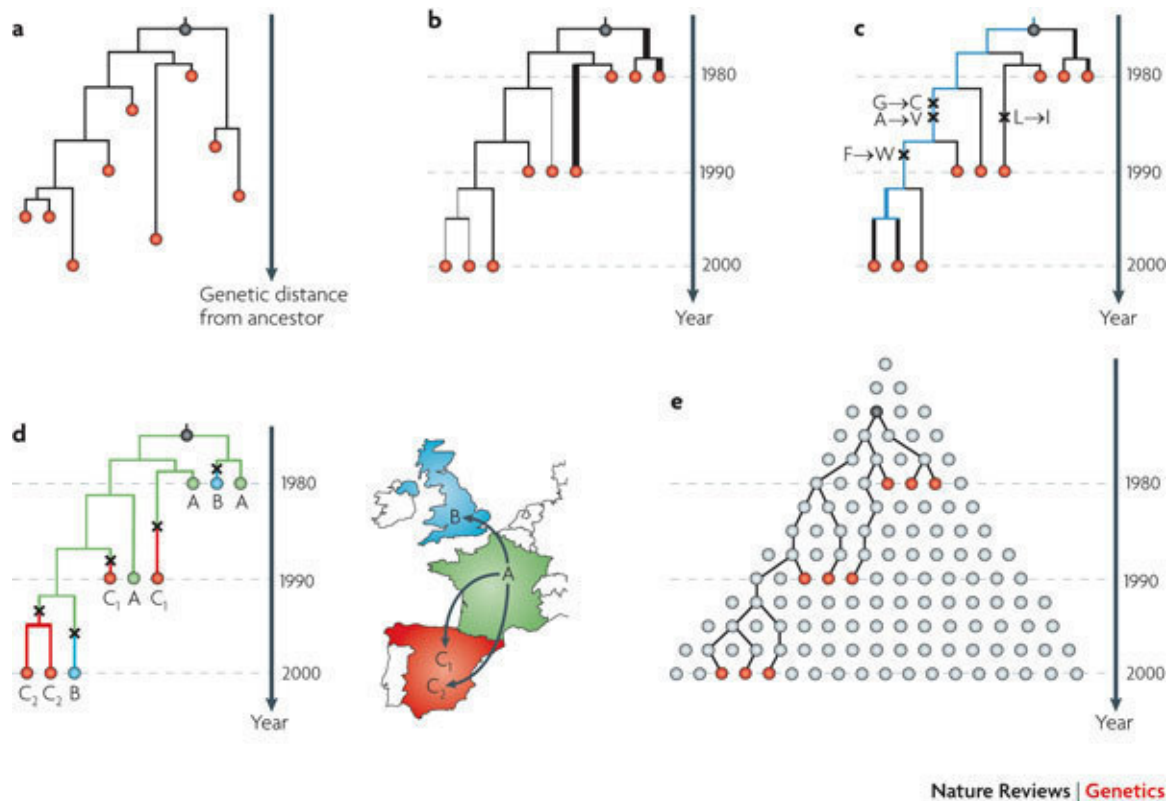
Les vaccins à virus inactivé ont été développés plus rapidement en raison des progrès récents dans les technologies de production (par exemple les cultures cellulaires sans sérum). Ils se caractérisent par un degré élevé d'immunogénicité, une bonne stabilité et une sûreté déjà éprouvée. La virulence de la souche virale est inactivée par un traitement au formaldéhyde ou à la chaleur. Cinq vaccins à virus inactivé sont en développement [Tableau 5] : trois projets chinois sont basés sur un virus C4, un projet taiwanais utilise un virus B4 et le dernier est développé avec une souche B3 à Singapour ; les trois premiers projets sont au stade III des essais cliniques (Li *et al*, 2014). Une préparation vaccinale avec une souche C4 conjuguée à un adjuvant d'alun présente une efficacité protectrice de 90% chez 10 000 participants dans un essai clinique de phase 3. De nombreux points restent à étudier : l'efficacité du vaccin contre des souches de sous génogroupes différents et les effets secondaires à long terme (Zhu *et al*, 2013). Les deux autres vaccins basés sur des souches de sous-génogroupes B3 et B4 sont encore en phase 1 (Kung *et al*, 2014).

Ces trois préparations présentent une bonne efficacité et pourraient constituer les premiers vaccins disponibles pour la population chinoise particulièrement affectée par les épidémies à EV-A71. Les améliorations à réaliser portent notamment sur le contrôle qualité de la production (il n'existe pas de standards internationaux concernant les lignées cellulaires, les souches virales et les antisérums). Concernant l'efficacité d'un vaccin contre l'EV-A71, la difficulté majeure est liée à la variabilité génétique de ce virus au cours du temps ; par conséquent, des études d'épidémiologie moléculaire seront impératives pour analyser l'évolution génétique des souches qui continueront à circuler lorsque le vaccin sera utilisé. Enfin, les vaccins en développement ne protègent pas des infections causées par les autres sérotypes d'entérovirus responsables de la maladie pied-main-bouche, CV-A16 notamment, qui sont en co-circulation avec l'EV-A71. Des approches de vaccins bivalents (EV-A71 / CV-A16) ou multivalents sont actuellement envisagées (Cai *et al*, 2014). Le développement des

vaccins anti-maladie pied-main-bouche pose la question de la rationalité de cette approche dans la mesure où les sérotypes d'entérovirus qui peuvent causer cette maladie sont très nombreux.



**LES METHODES  
D'ANALYSE  
PHYLODYNAMIQUE :  
CONCEPTS  
GENERAUX ET  
APPLICATION A  
L'EPIDEMIOLOGIE  
VIRALE**



Nature Reviews | Genetics

### Figure 30. Techniques d'étude phylodynamique d'une population virale

**a.** Représentation phylogénétique des relations de parentés génétiques dans un échantillon d'individus. La longueur des branches est exprimée en distance génétique par rapport à l'ancêtre commun.

**b.** Chronogramme construit en utilisant une horloge moléculaire estimée avec les dates d'isolement des souches, permettant la datation moléculaire des ancêtres communs établis par la reconstruction phylogénétique. La longueur des branches est exprimée en distance génétique par rapport à l'ancêtre commun à l'ensemble de l'échantillon en fonction de l'unité de temps utilisée.

**c.** Représentation phylogénétique en fonction du temps (datation moléculaire) montrant les changements en acides aminés survenus entre les ancêtres communs (adaptation moléculaire). La longueur des branches est exprimée en distance génétique par rapport à l'ancêtre commun à l'ensemble de l'échantillon en fonction de l'unité de temps utilisée.

**d.** Représentation phylogéographique montrant les relations phylogénétiques en fonction du temps et de la localisation géographique estimée pour chaque ancêtre commun. La probabilité de la localisation géographique est estimée d'après celles des souches de l'échantillon. Cette analyse permet de reconstruire l'histoire de la dissémination de l'agent pathogène.

**e.** Schéma représentant le principe de la théorie de la coalescence utilisée en phylogénie moléculaire. Chaque cercle représente un cas d'infection, positionné en fonction du temps. Les cercles rouges et gris indiquent respectivement les infections échantillonnées et celles qui ne l'ont pas été ; l'ancêtre commun à l'ensemble de l'échantillon est indiqué en noir.

(D'après Pybus *et al*, 2009).

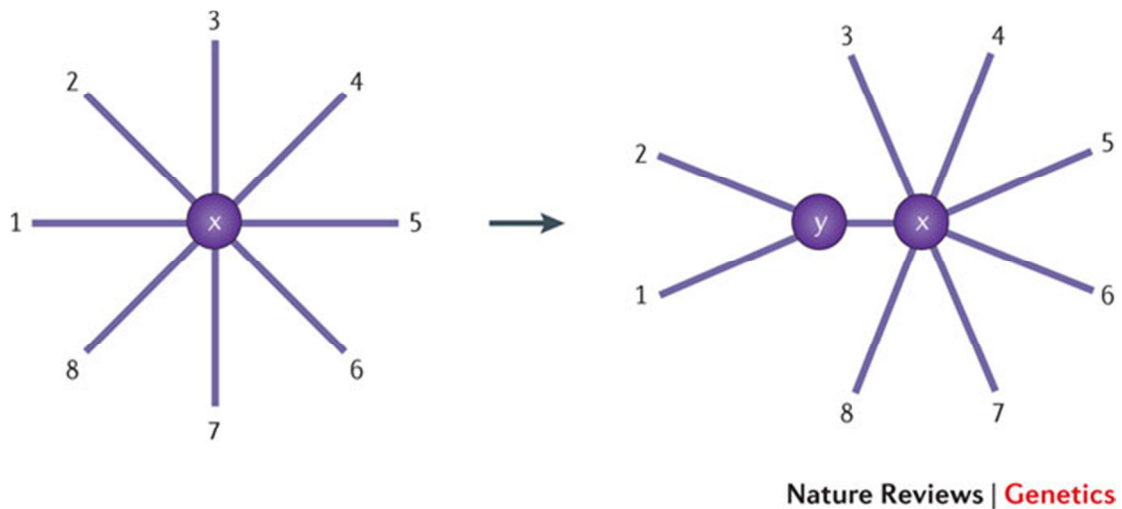
L'analyse phylodynamique désigne une discipline intégrative née de la convergence des méthodes utilisées en biologie évolutive et en épidémiologie : elle relie l'inférence\* phylogénétique et la dynamique des populations d'agents infectieux. Les développements les plus récents tentent d'intégrer à cet ensemble les modèles d'analyse utilisés en épidémiologie des maladies infectieuses (cet aspect ne sera pas traité). L'analyse phylodynamique est particulièrement adaptée à l'étude épidémiologique des virus à ARN car la vitesse de leur évolution est telle que les changements adaptatifs et les facteurs épidémiologiques qui en sont la cause peuvent être modélisés sur la même échelle de temps.

L'analyse phylodynamique appliquée à l'épidémiologie virale utilise le polymorphisme génétique pour modéliser conjointement la structure des populations virales en fonction du temps, leur démographie, leur dissémination entre différentes zones géographiques et la sélection (adaptation moléculaire) [Figure 30]. L'étude des processus phylodynamiques des virus est en plein essor grâce au séquençage des génomes, aux ressources informatiques disponibles pour l'analyse de jeux de données de plusieurs milliers de séquences et aux méthodes statistiques de modélisation des processus évolutifs et épidémiologiques. Au cours des dix dernières années, des progrès importants ont été réalisés pour intégrer les modèles\* utilisés en phylogénie moléculaire (en particulier la coalescence pour la datation moléculaire) à ceux qui sont utilisés pour analyser la démographie et la distribution géographique (phylogéographie).

## **I. Phylogénie moléculaire**

L'idée de s'intéresser aux relations de parenté génétique entre des organismes est issue de la théorie de l'évolution de Charles Darwin en 1859 publiée dans son ouvrage « The origin of species ». Les organismes vivants forment des populations appelées espèces qui peuvent subir des variations au cours du temps, permettant une adaptation à l'environnement, car les variations avantageuses sont conservées au fil des générations. Jusqu'aux années 1960, l'analyse comparée de la morphologie, du comportement et de la répartition géographique des espèces étaient les caractères phénotypiques observables pour construire les classifications des individus et des espèces.

La découverte de variations de séquences protéiques et nucléiques a ouvert un champ d'application considérable pour reconstruire l'histoire évolutive des individus ou celle des espèces avec l'analyse phylogénétique. La phylogénie se base sur l'identification de caractères homologues entre plusieurs individus ou taxons\*. L'arbre phylogénétique d'un



**Figure 31. Principe de l’algorithme « Neighbour Joining » (NJ)**

Le principe de la méthode NJ repose sur l’introduction d’une horloge moléculaire variable : les distances calculées entre deux taxons sont pondérées par la somme des distances entre chacun des taxons d’une paire considérée et tous les autres taxons de l’échantillon. L’algorithme NJ débute la construction de l’arbre phylogénétique par un graphe en étoile dans lequel le nœud central unique (x) est équidistant des taxons de l’échantillon. Les deux taxons (1 et 2) présentent la plus faible distance génétique dans l’échantillon des huit taxons, ce qui implique qu’ils partagent un ancêtre commun. Un arbre intermédiaire est reconstruit en reliant ces deux taxons par un nœud représentant leur ancêtre commun (y). Les branches (y,1) et (y,2) auront des longueurs différentes si les distances génétiques pondérées des taxons 1 et 2 sont différentes.

La matrice de distances est ensuite ré-estimée afin de prendre en compte la divergence moyenne de chaque paire de nœuds en fonction de la divergence moyenne de l’ensemble des nœuds. L’arbre est reconstruit en reliant les séquences les plus proches dans cette nouvelle matrice et en ajoutant un nœud terminal, formant leur ancêtre commun. Le processus est complet lorsque deux nœuds sont séparés par une seule branche.

(D’après Yang et Rannala, 2012).

groupe de taxons est inféré, en modélisant la répartition des caractères dans l'échantillon étudié. Différents types de caractères (structures morphologiques, voies biochimiques, gènes, séquences nucléotidiques, séquences protéiques) peuvent être utilisés pour reconstruire une phylogénie. Les phylogénies moléculaires traitent les caractères présents dans les séquences nucléiques ou protéiques. Une matrice quantitative est estimée afin de classer les états de chaque caractère, c'est-à-dire le type de nucléotide ou d'acide aminé présent à une position donnée chez chaque taxon. La reconstruction graphique de la phylogénie est réalisée en utilisant les données numériques de la matrice. L'arbre phylogénétique contient des nœuds (ou embranchements) connectés entre eux par des branches qui indiquent la persistance d'une lignée\* génétique au cours du temps.

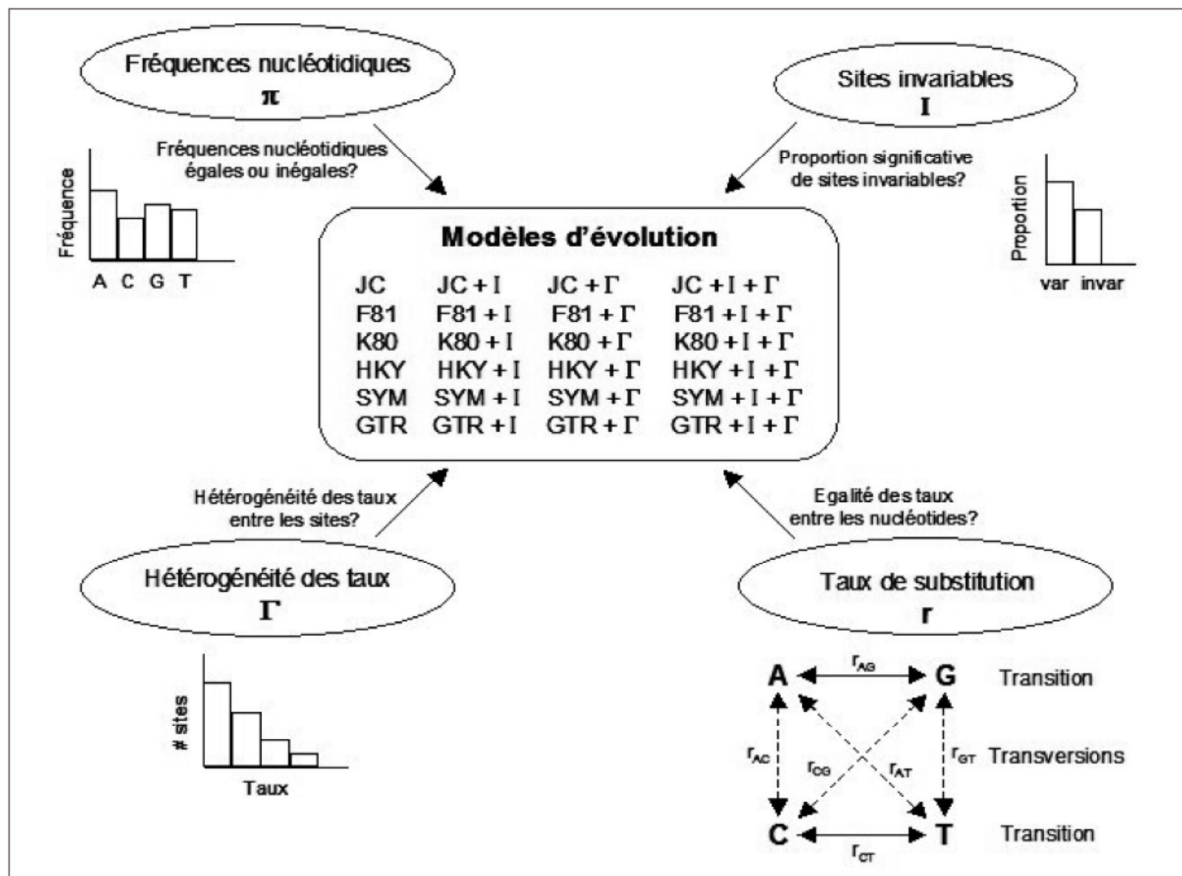
On distingue trois méthodes de reconstruction phylogénétiques, utilisant les séquences des gènes ou des protéines comme ensemble de caractères : (i) la méthode des distances, (ii) le maximum de parcimonie et (iii) les méthodes probabilistes utilisant la vraisemblance.

### **1) Méthode des distances**

La méthode dite des distances appliquée aux séquences nucléiques consiste à calculer une matrice de valeurs numériques qui représentent les distances génétiques entre chaque paire de taxons, c'est-à-dire les proportions de substitutions nucléotidiques. La méthode de calcul est dite paramétrique ; le processus des substitutions entre les séquences nucléotidiques peut être estimé à l'aide de nombreux modèles basés sur la combinaison de paramètres (cf. ci-dessous).

L'algorithme du « Neighbour Joining » utilise la matrice des distances pour construire le graphe qui représente l'arbre phylogénétique (Saitou et Nei, 1987). L'intérêt de cette méthode est de permettre la variation des taux d'évolution entre les lignées, donc l'intégration d'une horloge moléculaire. Le principe général de la méthode « Neighbour Joining » est exposé dans la figure 31. Cette technique de reconstruction a l'avantage d'être rapide et utilisable sur de grands jeux de séquences. Elle permet également d'identifier les lignées qui présentent des vitesses d'évolution différentes. Le degré de fiabilité de chaque embranchement peut être estimé par une méthode de ré-échantillonnage aléatoire, non paramétrique, appelée « bootstrap ». Les positions des sites de l'alignement sont tirées au sort pour créer un nouvel alignement. Un arbre est alors produit pour chaque répliquat et l'analyse est répétée un grand nombre de fois (1000 répliquats au minimum). Un arbre consensus indique la fréquence des regroupements des séquences et permet d'identifier des groupes





**Figure 32. Les paramètres principaux qui composent les modèles d'évolution moléculaire utilisés dans les méthodes probabilistes d'analyse phylogénétique**

Les modèles de substitution permettent de calculer la probabilité des changements entre les séquences d'un échantillon sous des hypothèses concernant leur évolution. Le remplacement d'un nucléotide par un autre est considéré comme un processus de Markov dans lequel la probabilité du changement de l'état  $i$  à l'état  $j$  à un site donné dépend uniquement de l'état  $i$ , pas des états précédents. Le processus est supposé homogène : cette probabilité ne change pas le long de chaque branche de l'arbre. Un modèle d'évolution moléculaire peut être représenté mathématiquement par une matrice 4 par 4 des taux instantanés de remplacement  $r$  de chaque nucléotide (A, T, C, G) par chaque nucléotide alternatif ( $r_{AG}$ ,  $r_{GT}$ ,  $r_{CT}$ ...), changement du nucléotide  $i$  par le nucléotide  $j$  dans un intervalle de temps infinitésimal. La matrice du modèle Jukes et Cantor le plus simple possède des taux instantanés identiques pour tous les changements. De nombreuses versions de la matrice ont été proposées jusqu'au modèle GTR (General Time Reversible) le plus complet car il comprend aussi des fréquences  $\pi$  propres à chaque nucléotide. Un modèle d'évolution moléculaire peut aussi comporter d'autres paramètres : les fréquences nucléotidiques  $\pi$ , la fraction des sites invariables  $I$  et l'hétérogénéité des taux de substitutions entre les sites modélisée par une fonction gamma (et représentée dans la fonction de vraisemblance par le paramètre alpha de la fonction gamma).

(D'après Delsuc et Douzery, 2004).

monophylétiques\*. Mais les informations obtenues grâce à la topologie\* issue d'une analyse par la méthode des distances sont peu nombreuses et dépendent du modèle d'évolution utilisé.

## 2) Maximum de parcimonie

Le maximum de parcimonie est une méthode statistique non paramétrique. Elle consiste à obtenir une topologie phylogénétique qui minimise le nombre de "pas" (mutations / substitutions) nécessaires pour relier une séquence à une autre.

## 3) Méthodes probabilistes - Vraisemblance

Les méthodes qui reposent sur un raisonnement probabiliste supposent que les événements évolutifs à l'échelle moléculaire obéissent à certaines lois de probabilité définies *a priori*. Elles nécessitent de définir un modèle d'évolution préalablement à l'analyse des caractères. Les paramètres du modèle permettent de décrire le processus d'évolution des séquences. Pour que cette description soit la plus réaliste possible, il incorpore en général les paramètres suivants : les fréquences nucléotidiques, les taux de substitutions, la fraction de sites invariables et l'hétérogénéité des taux de substitutions entre sites [Figure 32]. De nombreux modèles de complexité croissante ont été conçus ; ils présentent des relations hiérarchiques entre eux ; le plus général est le modèle « General Time Reversible » (GTR).

Avec un modèle d'évolution moléculaire, il devient possible d'exprimer la probabilité qu'un arbre évolutif aboutisse à l'observation de l'état de chaque caractère contenu dans les données : par exemple tel nucléotide (ou acide aminé) à telle position d'une séquence particulière. La fonction de vraisemblance permet de calculer la probabilité d'observer les données conditionnellement au modèle d'évolution moléculaire choisi *a priori*. Les valeurs numériques optimales des paramètres du modèle d'évolution moléculaire sont estimées lors d'une analyse itérative, grâce à l'algorithme du maximum de vraisemblance. Chaque jeu de données\* possède son propre ensemble de valeurs optimales, celles qui maximisent la vraisemblance des données observées. Le choix du modèle GTR pourrait sembler judicieux puisqu'il est le plus complet, mais son utilisation présente l'inconvénient d'allonger le temps de calcul et surtout le nombre des paramètres considérés accroît la variance de chacun d'eux. Des tests statistiques ont donc été développés pour comparer l'adéquation des modèles d'évolution moléculaire aux données : le « Likelihood Ratio Test » (LRT) et le critère d'information d'Akaike (AIC)\*.

Ensuite, à partir du modèle sélectionné *a priori*, un premier arbre phylogénétique est estimé au hasard parmi les nombreuses possibilités grâce à un algorithme heuristique : « Stepwise addition » (Felsenstein, 1981) ou BioNJ (Gascuel, 1997). Cette procédure consiste

à produire un arbre à partir de quelques taxons présents au sein de l'alignement (souvent les plus éloignés génétiquement, BioNJ) et à étoffer cet arbre au fur et à mesure en ajoutant de nouveaux taxons étape par étape. Une probabilité de vraisemblance sera estimée pour cet arbre de départ. Les étapes suivantes consistent à effectuer des réarrangements de cet arbre initial par différentes méthodes (Nearest Neighbor Interchange, NNI ; Sub-tree Pruning and Regrafting, SPR) afin de produire une grande quantité d'arbres phylogénétiques présentant des topologies différentes (longueurs des branches et topologie). Ensuite, les paramètres de la topologie obtenue sont comparés à ceux du modèle d'évolution moléculaire qui maximisent la vraisemblance des données. Le choix de l'arbre ayant la topologie optimale se fera sur la valeur de vraisemblance : la topologie ayant la plus forte probabilité de conduire aux données (Felsenstein, 1973). La procédure d'inférence phylogénétique par maximum de vraisemblance est implémentée dans plusieurs logiciels dont phyML (Guindon et Gascuel, 2003).

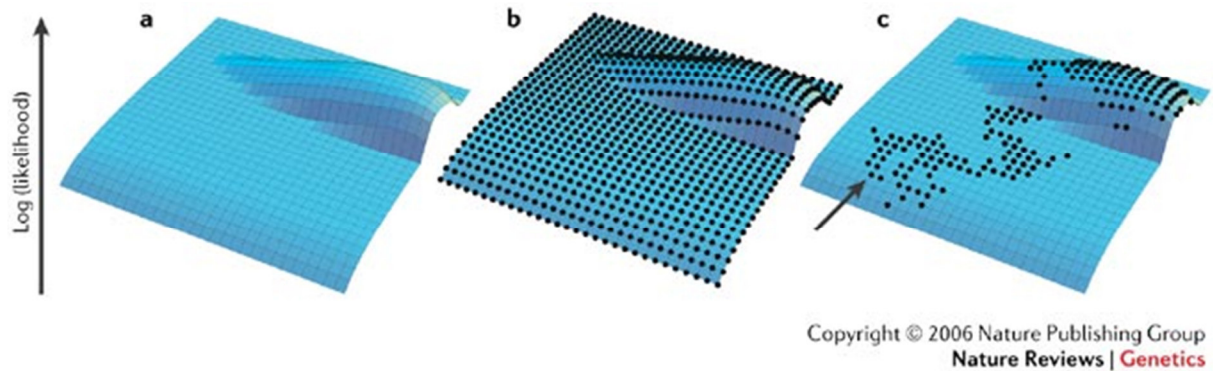
#### 4) Généralités sur l'inférence Bayésienne et son application en phylogénie

Le théorème de Bayes est une fonction mathématique de probabilité qui a été inventée en 1748 par un pasteur britannique, Thomas Bayes. Basée sur les lois du hasard et de l'incertitude, elle se définit par l'équation suivante «  $\Pr (A|B) = \Pr (B|A) \times \Pr (A) / \Pr (B)$  » et permet de calculer la façon dont évolue la probabilité d'un événement A, sachant qu'un autre événement B a eu lieu. Ce théorème présente un intérêt pour décrypter des phénomènes dont les paramètres sont corrélés. Appliqué à la phylogénie moléculaire, le théorème de Bayes permet de calculer des probabilités postérieures (calculées *a posteriori*) d'arbres phylogénétiques à partir des probabilités des paramètres d'un modèle défini *a priori*. La probabilité postérieure est la probabilité de l'hypothèse H (c'est-à-dire une topologie phylogénétique) sachant les données D (séquences),  $\Pr (H | D)$ , ce qui diffère du maximum de vraisemblance où l'on cherche la probabilité d'observer les données sous une hypothèse H,  $\Pr (D | H)$ . La probabilité postérieure d'une hypothèse se calcule par le théorème de Bayes selon l'équation suivante :

$$\Pr (H | D) = \Pr (D | H) \times \Pr (H) / P (D)$$

$\Pr (D | H)$  est la fonction de vraisemblance,  $\Pr (H)$ , la probabilité a priori de l'hypothèse H et  $\Pr (D)$ , la probabilité des données. On peut considérer la probabilité postérieure d'une hypothèse comme la probabilité que cette hypothèse soit vraie sachant les données.

L'approche bayésienne a été appliquée à la reconstruction des phylogénies en 1996 (Rannala et Yang, 1996) et s'est rapidement imposée dans ce domaine. L'approche statistique bayésienne en phylogénie consiste à combiner la probabilité *a priori* d'une topologie  $\Pr (T)$



**Figure 33. Technique de la Chaîne de Markov Monte Carlo (MCMC)**

La technique de la Chaîne de Markov Monte Carlo (MCMC) est utilisée pour estimer la distribution de probabilité postérieure d'un ensemble de paramètres sans avoir à explorer l'ensemble de l'espace des paramètres.

a. Représentation de la surface des vraisemblances d'une combinaison de deux paramètres. Le relief est proportionnel aux valeurs de vraisemblance.

b. Cas de l'exploration de l'intégralité de la surface des vraisemblances d'une combinaison de deux paramètres. Le relief est proportionnel aux valeurs de vraisemblance.

c. Exploration de la surface des vraisemblances d'une combinaison de deux paramètres à partir de la technique de la chaîne MCMC. Cette chaîne explore de manière dirigée l'espace des vraisemblances en ciblant les valeurs les plus élevées. Le début de l'exploration est représenté par la flèche noire.

**(D'après Excoffier et Heckel., 2006).**

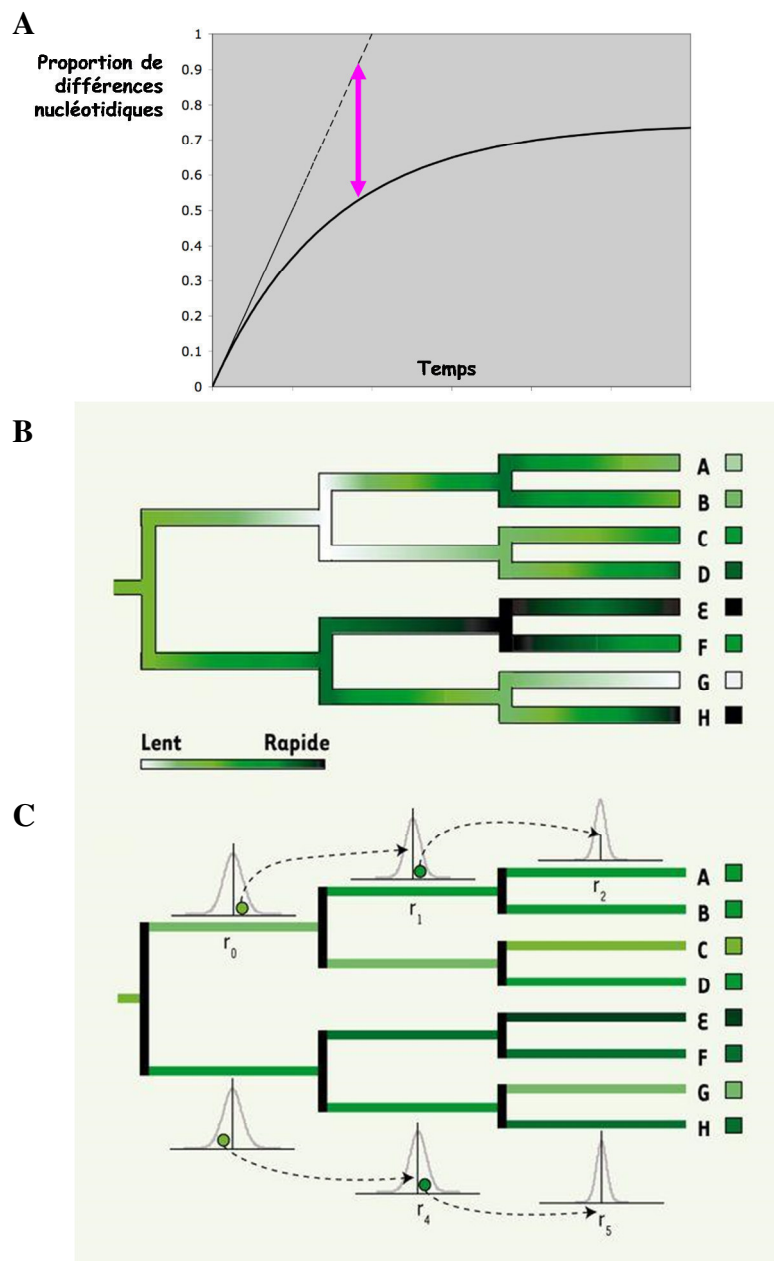
avec la vraisemblance  $\Pr(D | T)$  pour générer une distribution de probabilité postérieure  $\Pr(T | D)$  d'un ensemble de topologies en utilisant le théorème énoncé ci-dessus :

$$\Pr(T | D) = \Pr(D | T) \times \Pr(T) / P(D).$$

La probabilité postérieure d'un arbre particulier peut être interprétée comme la probabilité que cet arbre soit vrai sachant les données. L'intérêt de l'inférence bayésienne repose sur la possibilité d'estimer la distribution de la probabilité postérieure d'un échantillon d'arbres pour un jeu de données, l'arbre qui présente la probabilité postérieure la plus élevée est identifié pour représenter la phylogénie.

Le calcul exhaustif des probabilités postérieures devient impossible pour l'ensemble des topologies possibles pour un jeu de données de plus de quelques dizaines de séquences. Des méthodes de statistique numérique appelées chaînes de Markov et de Monte Carlo (MCMC) ont dû être intégrées à l'analyse bayésienne pour calculer ces probabilités. Cette chaîne est une sorte de « marche guidée » à travers l'espace multidimensionnel des paramètres (topologie, longueur des branches de l'arbre, paramètres du modèle d'évolution moléculaire) qui permet l'estimation de la distribution de probabilité en échantillonnant les valeurs de ces paramètres de manière périodique [Figure 33]. L'exactitude de l'approximation de la distribution dépend du nombre de cycles (ou générations) de la chaîne. Chaque génération implique une modification aléatoire de la topologie, de la longueur des branches de l'arbre et des valeurs des paramètres du modèle d'évolution moléculaire. L'estimation de la distribution de probabilité postérieure des arbres phylogénétiques est effectuée avec l'algorithme de Metropolis-Hastings (Metropolis *et al*, 1953 ; Hastings, 1970).

Pratiquement, la première génération de la chaîne MCMC utilise un arbre produit au hasard à partir des séquences et la probabilité postérieure de cet arbre est estimée. La chaîne réalise une nouvelle génération à partir d'un arbre comportant un nouvel état (modification aléatoire de la topologie et/ou des longueurs des branches et/ou des paramètres du modèle d'évolution moléculaire). La probabilité postérieure de cet arbre est estimée (probabilité que l'arbre soit vrai sachant les données). Le nouvel état est conservé s'il est meilleur que le précédent (probabilité postérieure supérieure à celle de l'autre topologie). Ainsi, au fil des générations, les estimations de chaque paramètre convergent vers la valeur de probabilité la plus élevée aboutissant à la phylogénie la plus probable. Le nombre de générations nécessaires pour atteindre cette convergence est appelé la période d'allumage (« burning »). Par sa capacité d'analyse de jeux de séquence de taille élevée, l'approche bayésienne permet de s'affranchir des limites de la phylogénie moléculaire en modélisant d'autres aspects que



**Figure 34. Horloges moléculaires et datation**

**A. Représentation de l'hypothèse de l'horloge moléculaire.** Les pointillés représentent la proportion attendue si l'évolution suit un processus constant (horloge moléculaire). Le trait plein représente la proportion de différences nucléotidiques la plus souvent observée.

**B. Variations graduelles des taux d'évolution le long des branches d'une phylogénie.** Les taux d'évolution varient sur chaque branche de la représentation phylogénétique de 8 taxons, définissant différentes trajectoires évolutives.

**C. Modèle d'horloge moléculaire assouplie avec autocorrélation des taux.** Les variations graduelles de taux d'évolution sont modélisées selon une distribution normale. Chaque branche est caractérisée par un taux unique, qui est la moyenne de son taux initial et de son taux final. Le taux  $r_1$  d'une branche descendante est tiré (disque vert) dans une distribution normale centrée sur le taux  $r_0$  de la branche ascendante. Le taux  $r_1$  est ici plus rapide que  $r_0$ . Les autres relations d'ordre illustrées sont :  $r_2 > r_1 > r_0$ ,  $r_3 > r_4$ , et  $r_5 > r_4$ . Une variance de cette distribution normale des taux respectivement importante ou faible entraîne un écart important ou faible par rapport à l'hypothèse d'horloge moléculaire.

(D'après Douzery *et al.*, 2006)

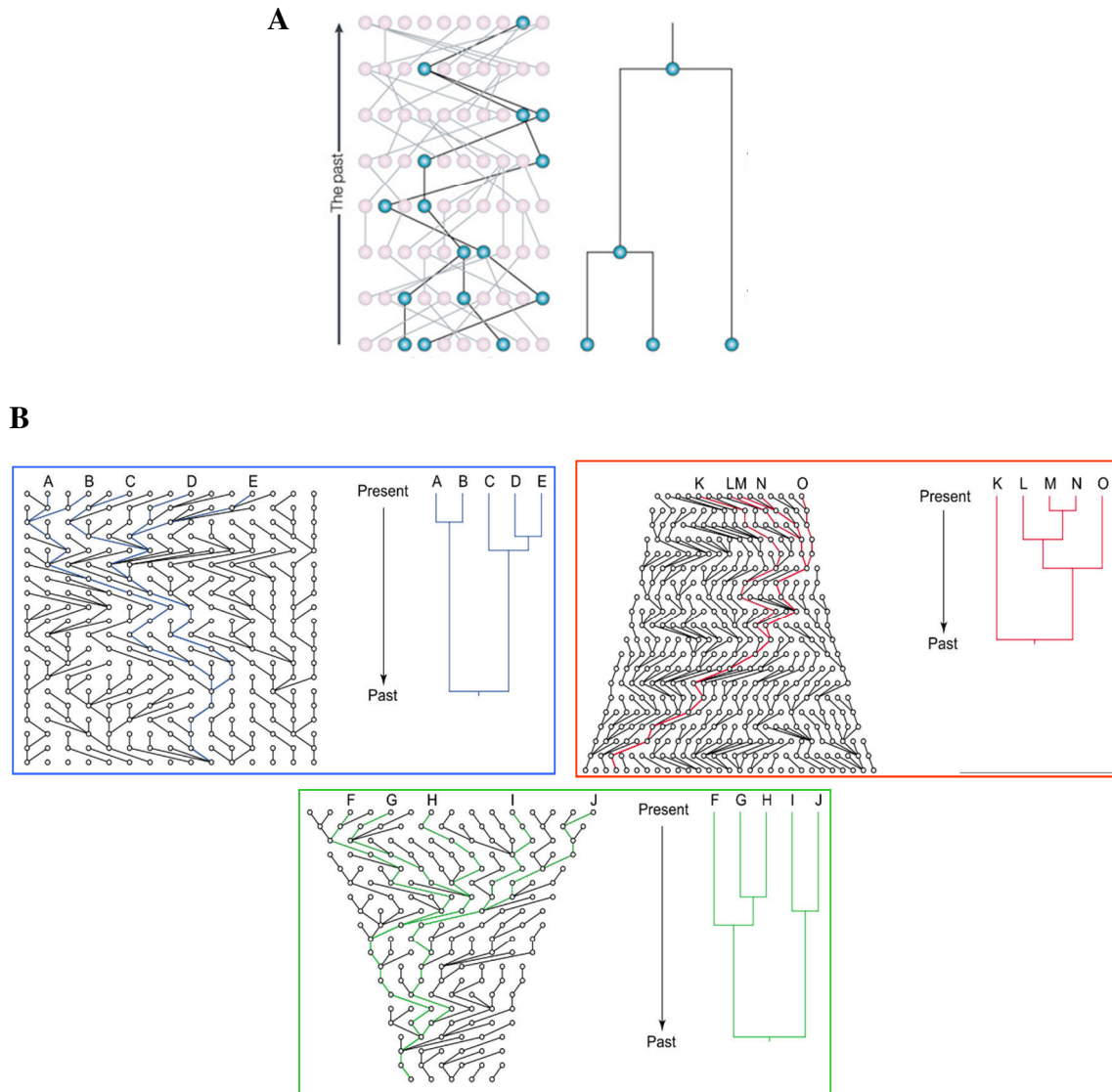
l'évolution des séquences et en intégrant les notions de datation moléculaire, de démographie des populations virales et de déplacement/dissémination géographique.

L'implémentation de ces algorithmes dans des programmes tels MrBayes ou BEAST (voir paragraphe VI, ci-dessous) permet d'analyser des jeux de données conséquents et confère à cette approche un intérêt indéniable.

## II. Datation moléculaire

L'invention du concept d'horloge moléculaire en 1965 par Emile Zuckerkandl et Linus Pauling a été un point fondamental du développement de la phylodynamique (Douzery *et al*, 2006). En effet, il établit que dans le cas « idéal », les génomes accumulent des mutations selon un processus uniforme et constant dans le temps (horloge moléculaire\* globale) [Figure 34A] (Gojobori *et al*, 1990). Sans ce concept, il serait impossible de coupler une quantité d'évolution moléculaire parmi plusieurs taxons (représentée par un arbre phylogénétique) avec des points de calibration temporelle, pour réaliser la datation moléculaire\*. Avec l'horloge moléculaire, il est possible d'estimer l'âge d'apparition des espèces pour lesquelles aucune information paléontologique n'est disponible.

La constance de l'horloge moléculaire est rarement vérifiée chez les micro-organismes, en raison de l'existence de processus génétiques (mutations réverses / séquentielles et recombinaison) qui font varier les taux d'évolution entre les lignées [Figure 34B] (Douzery *et al*, 2006). De nouvelles hypothèses d'horloge moléculaire dites « assouplies » ont été formulées (Thorne *et al*, 1998). Elles se basent sur l'héritabilité du taux d'évolution lors de la spéciation de deux taxons, autrement dit, elles modélisent les variations des taux d'évolution entre les taxons selon une distribution mathématique. Des différences graduelles de taux d'évolution peuvent être estimées le long des branches de la représentation phylogénétique [Figure 34C]. Les méthodes d'horloge moléculaire assouplie ont été appliquées à la phylogénie en l'intégrant à un cadre statistique Bayésien et à une chaîne MCMC. L'analyse phylogénétique combine à la fois des données moléculaires et temporelles, permettant l'estimation des âges absolus estimés de divergence entre des taxons (Drummond *et al*, 2006).



**Figure 35. Généalogie, coalescence et phylogénie**

**A. Principe de la coalescence.** La figure de gauche représente la généalogie d'une population d'étude en fonction du temps. Les lignes noires et les cercles bleus représentent les individus échantillonnés et leurs ancêtres communs. La figure de droite représente la représentation phylogénétique découlant de la généalogie. (D'après Rosenberg et Nordborg, 2002)

**B. Influence de la variation de la démographie sur l'inférence phylogénétique.** La figure illustre la généalogie d'échantillons d'individus ayant différents scénarios démographiques en fonction du temps : population d'effectif constant (bleu), population soumise à une décroissance exponentielle (rouge) et population soumise à une croissance exponentielle (vert). Pour chaque scénario, cinq séquences ont été échantillonnées. Les généalogies estimées à droite montrent l'influence de la démographie de la population entière sur la représentation phylogénétique reconstruite avec l'échantillon. A partir des méthodes basées sur la coalescence, il est possible d'inférer l'histoire des changements de taille de population à partir d'un échantillon de séquences géniques obtenues d'une population. (D'après Emerson *et al*, 2001)



### III. Coalescence et démographie

Les variations de la démographie\* d'une population dans le temps, l'histoire démographique, laissent des signatures génétiques au sein de la descendance. Retracer cette histoire permet de comprendre certains processus évolutifs impliqués dans la transmission et la dissémination des virus. Des progrès notables dans la compréhension de la génétique des populations virales ont été réalisés en utilisant notamment le modèle de coalescence\* (Kingman, 1982). Ce modèle stochastique décrit la relation qui existe entre l'histoire démographique d'une population de grand effectif et la généalogie\* (représentée par un arbre phylogénétique) que partagent des individus échantillonnés au hasard dans cette population [Figure 35A]. Les séquences moléculaires issues d'un échantillonnage dans le temps peuvent être utilisées pour reconstruire l'histoire démographique des populations naturelles. La figure 35B illustre la généalogie d'échantillons d'individus tirés de trois populations qui présentent différents types démographiques (population évoluant à effectif constant, population en décroissance exponentielle, population en croissance exponentielle).

Le modèle de coalescence fournit la distribution de probabilités des périodes entre deux événements de coalescence dans la généalogie d'un échantillon, distribution qui dépend de la fonction représentant le type démographique. La fonction démographique peut être estimée à partir de données moléculaires en combinant, dans une approche bayésienne, les fonctions de vraisemblance du modèle phylogénétique et du modèle de coalescence. Les modèles démographiques sont des fonctions mathématiques caractérisée par des paramètres (Pybus et Rambaut, 2002) : le modèle d'effectif constant comprend un seul paramètre, le modèle de croissance exponentielle (taux de croissance constant au cours du temps) a deux paramètres, le modèle de croissance logistique (taux de croissance décroissant au cours du temps) a trois paramètres, le modèle de croissance en expansion (augmentation du taux de croissance au cours du temps) a trois paramètres.

Les organismes des populations naturelles présentent des histoires démographiques bien plus complexes que celles décrites par les modèles paramétriques qui représentent des populations « idéales ». Des modèles démographiques non paramétriques « flexibles » ont été développés. Le modèle appelé « Skyline Plot » permet une estimation de l'histoire de la taille de la population à partir de sa généalogie en s'affranchissant des limites imposées par les modèles paramétriques. Le principe du modèle Skyline Plot est d'estimer la généalogie de la population et de la décomposer en plusieurs parties (intervalles de coalescence) par la méthode mathématique de décomposition séquentielle appelée « piecewise-constant ». Un

paramètre dit de diversité génétique est estimé pour chaque intervalle de coalescence (Pybus *et al*, 2000). Ce paramètre est proportionnel au temps de génération et à la taille efficace de la population (l'effectif d'une population idéale présentant tous les changements génétiques qui ont lieu dans l'intégralité de la population réelle). Les variations des valeurs numériques de ce paramètre permettent d'obtenir une approximation de la démographie de la population réelle. L'estimation des paramètres du modèle « Skyline Plot » peut être effectuée par une méthode Bayésiennes, méthode appelée « Bayesian Skyline Plot ; BSP » (Drummond *et al*, 2005). Cette approche permet d'estimer une distribution postérieure de la taille de la population effective au cours du temps suivant une combinaison de modèles d'évolution moléculaire et d'horloge moléculaire grâce à des itérations MCMC.

La méthode BSP permet donc d'estimer l'évolution de la diversité génétique au cours du temps, ce qui donne une représentation des dynamiques démographiques ancestrales d'une population d'étude à partir de données moléculaires. Toutefois, le nombre d'intervalle de coalescence doit être déterminé *a priori*.

#### **IV. Phylogéographie**

La phylogéographie concerne l'étude des principes et des processus liés à la distribution géographique des lignées qui constituent des populations proches génétiquement : elle est donc centrée sur l'arbre phylogénétique intraspécifique. La diffusion spatiale des agents pathogènes comme celle des autres organismes laisse des empreintes mesurables au niveau des séquences génétiques (Holmes, 2004). Les phylogénies des virus dont l'évolution se déroule sur le même tempo que les processus écologiques qui façonnent leur diversité, contiennent une information sur la dynamique spatiale en plus des dynamiques temporelle et démographique. Pour les virus qui ont une multiplication strictement humaine, la dynamique spatiale est conditionnée par les changements qui interviennent dans la population hôte, les variations de la densité des individus sensibles et les déplacements des individus infectés à l'intérieur d'une région géographique ou entre des régions distantes. C'est ainsi que la diffusion spatiale peut être retracée à l'aide de modèles en utilisant les séquences moléculaires, les signatures génétiques propres à chaque population de souches en circulation.

En retraçant ces processus de diffusion spatiale, on peut élucider des facteurs clés de leur transmission à petite et grande échelle géographique et améliorer la compréhension de l'épidémiologie virale, autrement dit de l'écologie virale. Les méthodes phylogénétiques

utilisant la coalescence et des informations temporelles et spatiales des souches virales étudiées permettent de caractériser le profil de transmission d'un virus donné grâce un cadre statistique Bayésien couplé à une chaîne MCMC. C'est ainsi que la généalogie qui résultera de cette analyse contiendra des informations évolutives, de datation, de démographie et d'origines géographiques probables de chaque cluster de souches.

Ce processus qui permet d'associer des caractères discrets (origine géographique) à l'inférence phylogénétique est appelé « Continuous-time Markov Chain » (Lemey *et al*, 2009). Il va estimer une distribution de probabilités de l'origine géographique probable de chaque nœud au sein de la généalogie, permettant d'évaluer le taux de diffusion de la population virale. On peut améliorer les estimations en couplant l'analyse à un algorithme « Bayesian Stochastic Search Variable Selection ; BSSVS », qui permet d'accroître considérablement l'efficacité statistique en n'estimant que les taux de migration ayant des probabilités non nulles (Lemey *et al*, 2009).

## **V. Evolution adaptative des protéines**

Durant leur évolution, les organismes peuvent subir des mutations synonymes ou non synonymes, à l'origine de leur diversité génétique ou protéique. Ces mutations peuvent conférer des avantages sélectifs pour la population. L'approche de reconstruction des états ancestraux (Ancestral State Reconstruction, ASR) a été développée durant les années 60 afin d'inférer une séquence ancestrale d'hémoglobine de mammifères et par la suite synthétiser cette protéine ancestrale pour en étudier les propriétés biochimiques, ce qui fut possible durant les années 90 (Pauling et Zuckerkandl, 1963).

Différentes méthodes ASR ont été développées au cours des dernières années en se basant sur des approches par maximum de parcimonie, par maximum de vraisemblance et par inférence Bayésienne. Les différentes hypothèses introduites dans chaque méthode font qu'elles sont adaptées à des échelles de temps différents et donc à des utilisations différentes. Les approches par maximum de vraisemblance et par inférence Bayésienne ont l'avantage d'incorporer les modèles d'évolution moléculaire dans les inférences ASR réalisées, ce qui les rend plus adaptées à l'étude de l'histoire évolutive des virus. La méthodologie ASR par inférence Bayésienne est encore en développement et nécessite des ressources importantes pour le calcul. L'approche par maximum de vraisemblance est donc appropriée à l'évolution des virus. Elle a l'avantage de pouvoir inférer les séquences ancestrales à chaque nœud

interne d'une phylogénie (Yang *et al*, 1995). Elle permet d'estimer les substitutions synonymes et non synonymes le long des branches d'une phylogénie afin d'identifier les cas d'adaptation moléculaire. Elle nécessite trois éléments : un alignement nucléotidique ou protéique, un arbre phylogénétique représentant les données et un modèle d'évolution moléculaire.

La plateforme web serveur Fast-ML permet de réaliser une analyse ASR à partir d'un alignement et d'un arbre phylogénétique lui correspondant. Elle emploie une approche d'ASR par maximum de vraisemblance et un modèle d'évolution moléculaire. Deux algorithmes sont utilisés conjointement : un premier qui prend en compte les insertions/délétions potentiellement présentes dans l'alignement et un second qui va inférer l'état ancestral le plus probable d'une position nucléotidique qui n'est pas un gap. Un test statistique par probabilité postérieure est également implémenté afin de prendre en compte la fiabilité de chaque inférence d'état ancestral.

## **VI. Le logiciel BEAST**

Le logiciel BEAST (Bayesian Evolutionary Analysis Sampling Trees ; Drummond et Rambaut, 2007) est un programme qui implémente des méthodes statistiques bayésiennes ainsi que l'algorithme Metropolis-Hasting MCMC afin de réaliser des inférences phylogénétiques des séquences moléculaires. Il permet des combinaisons de modèles d'évolution moléculaire, d'horloge moléculaire, des modèles démographiques ainsi que d'autres modèles (phylogéographie, ASR, ...). BEAST constitue une plateforme d'analyse de l'histoire évolutive d'une population d'étude en utilisant un échantillon de séquences moléculaires. Il bénéficie de mises à jour régulières pour intégrer de nouveaux modèles et s'adapter à l'évolution de la puissance de calcul des machines informatiques (processeurs multi cœurs par exemple).

### **1) Estimation de généalogies les plus probables (taux d'évolution et TMRCA)**

Par rapport à d'autres outils d'inférence phylogénétique par méthodes bayésiennes, le logiciel BEAST est conçu pour estimer des généalogies enracinées à partir d'un jeu de données et de points de calibration temporelle propres à chaque séquence (les dates de détection des souches). Ainsi, l'analyse phylogénétique en utilisant BEAST consiste à estimer un grand nombre de généalogies pour les données moléculaires considérées (une topologie à

chaque génération de la chaîne de Markov). Les valeurs numériques calculées pour chaque paramètre sont optimisées jusqu'à ce que les estimations convergent sur une valeur. La progression de la chaîne de Markov est suivie grâce à un programme compagnon de BEAST, le logiciel « Tracer ». Lorsque l'estimation de tous les paramètres du modèle a convergé vers des valeurs numériques maximum, le programme « Treeannotator » permet de rechercher la généalogie qui présente la vraisemblance la plus élevée parmi l'échantillon d'arbres sélectionnés au cours de la chaîne de Markov.

L'arbre le plus probable est associé à un ensemble d'informations (valeurs numériques des paramètres estimés pendant l'analyse). Dans cet ensemble figure en particulier la date d'émergence de l'ancêtre commun à tous les taxons (« time to the most recent common ancestor » ou TMRCA) estimée d'après le modèle d'horloge moléculaire choisi. Une valeur TMRCA est aussi associée à chaque nœud dont la densité de probabilité postérieure est supérieure à un seuil fiable (en général 0,9). L'échantillonnage d'un grand nombre d'arbres par la chaîne de Markov rend possible le calcul d'un intervalle de crédibilité ou « highest posterior density interval, HPD ». Cet intervalle peut être assimilé à un intervalle de confiance à un seuil statistique choisi. L'intervalle "95% HPD" d'un paramètre donné contient 95% des valeurs échantillonnées par la chaîne de Markov pour ce paramètre.

## **2) Estimation de la démographie**

Le logiciel BEAST permet l'inférence des dynamiques démographiques de la population échantillonnée simultanément à celle de la généalogie. Les algorithmes de plusieurs modèles démographiques basés sur la coalescence, sont implémentés. Le logiciel permet de tester statistiquement des scénarios démographiques préétablis sur un jeu de données : évolution d'une population de taille constante, en croissance exponentielle, en croissance logistique et en expansion. BEAST intègre également un modèle démographique plus flexible « Bayesian Skyline Plot » (BSP) qui permet de représenter graphiquement les variations démographiques de la population d'étude lorsque son effectif varie en fonction du temps (succession de phases de croissance et de décroissance). Ce profil BSP peut être reconstruit à partir du logiciel Tracer.

## **3) Estimation de la distribution géographique des lignées**

Le cadre statistique bayésien de BEAST ouvre la possibilité de reconstruire les états ancestraux à partir d'hypothèses évolutives et d'états discrets. Si ces derniers sont des informations géographiques relatives à l'échantillon des séquences analysées, l'analyse

phylogénétique peut inclure l'estimation de la probabilité de l'origine géographique de chaque nœud (ancêtre commun) de la généalogie. Les taux de migration entre les différents nœuds sont également estimés et peuvent être utilisés pour obtenir un profil géographique retraçant les flux de transmission des organismes étudiés entre les zones géographiques représentées dans l'échantillon initial.



**OBJECTIFS DES  
TRAVAUX DE LA  
THESE**



Les travaux réalisés pendant cette thèse portent sur l'EV-A71 et les sérotypes apparentés principalement responsables de la maladie pied-main-bouche. Les objectifs sont les suivants.

### **1) Epidémiologie moléculaire de l'EV-A71 et des principaux EV-A responsables du syndrome pied-main-bouche**

Analyser la dynamique épidémiologique de ce virus à partir de données moléculaire permettrait d'identifier les facteurs viraux liés à la transmission. Quelques sérotypes d'EV-A ont fait l'objet d'étude d'épidémiologie moléculaire dans le but de décrire une épidémie particulière ou la diversité des souches au cours d'une période définie. Dans cette première partie, notre objectif est de comparer les principaux sérotypes circulants d'EV-A (EV-A71, CV-A16, CV-A6, CV-A10) sous les aspects de leur diversité génétique (définition des génogroupes et sous génogroupes), leur histoire évolutive (répartition des populations virales au cours du temps) et de leur évolution démographique. Pour cette étude, nous utilisons des méthodes de modélisation en phylogénie moléculaire permettant l'analyse des jeux de séquences de grande taille. De plus, nous cherchons à estimer la dynamique migratoire des différents sous génogroupes de l'EV-A71 entre l'Europe et l'Asie (flux de transmissions des populations virales entre différentes zones géographiques).

### **2) Evolution structurale de la capsid de l'EV-A71**

Dans la deuxième partie de la thèse, nous cherchons à analyser l'évolution moléculaire de la région génomique P1 codant les protéines de capsid de l'EV-A71, en identifiant les mutations qui surviennent pendant la circulation des souches dans la population hôte. Ces changements en acides aminés peuvent être favorables pour le virus et constituer les déterminants d'une évolution adaptative, par exemple vis-à-vis des réponses immunitaires de l'hôte.

### **3) Rôle de la recombinaison génétique dans l'évolution de l'espèce EV-A**

Les souches d'EV-A71 circulent en même temps que celles d'autres sérotypes d'EV-A. Ceci constitue un facteur épidémiologique nécessaire à la survenue de co-infections favorisant la recombinaison génétique. Ce processus est un facteur majeur d'évolution et de diversité des populations virales mais son impact sur l'évolution des EV-A demeure mal connu. L'objectif général de l'étude est de déterminer si la recombinaison est un facteur important dans

l'émergence des populations virale et des virus variants de l'EV-A71 et d'autres EV-A apparentés.

#### **4) Etude du pouvoir pathogène des EV-A avec un modèle cellulaire de barrière endothéliale microvasculaire de derme**

Les manifestations cutanées du syndrome pied-main-bouche sont bénignes mais elles participent à la transmission du virus. Il existe peu de modèle cellulaire approprié pour étudier le pouvoir pathogène des EV-A, en particulier la physiopathologie des syndromes PMB. L'objectif principal de cette dernière partie est de tester une lignée cellulaire d'endothélium microvasculaire appelée « HMEC-1 » (cellules endothéliales de derme humain) pour déterminer si elle peut être utilisée comme un modèle d'étude des mécanismes de franchissement des barrières endothéliales du derme par l'EV-A71 et les EV-A apparentés.



# **MATERIELS ET METHODES**

**Tableau 6.** Origines géographiques des souches génotypées au laboratoire en fonction de leurs sérotypes

Sérotipe	Origine géographique					
	France	Allemagne	Finlande	Autriche	Hongrie	Tunisie
<b>CV-A2</b>	17	1				
<b>CV-A4</b>	30					3
<b>CV-A5</b>	8	1				
<b>CV-A6</b>	115	4	6	1		
<b>CV-A8</b>	13					
<b>CV-A10</b>	31	3	4			
<b>CV-A14</b>	4		1	1		
<b>CV-A16</b>	141	8		3	6	
<b>EV-A71 B4/B5</b>	1					
<b>EV-A71 C1</b>		1				
<b>EV-A71 C2</b>	42	11			2	
<b>EV-A71 C4</b>	2				6	

## **I. Souches d'entérovirus**

### **1) Echantillon constitué pour étudier l'épidémiologie moléculaire des EV-A**

Les souches d'EV-A utilisées pour cette étude ont été obtenues dans le cadre de la surveillance des infections par des entérovirus (incluant la surveillance de la poliomyélite) effectuée dans différents pays européens : France (n=404), Fédération de Russie (n=30), Allemagne (n=29), Hongrie (n=14), Autriche (n=5), Finlande (n=11), Azerbaïdjan (n=1), Kazakhstan (n=1), Kirghizistan (n=1) et Tunisie (n=3). Ces souches proviennent de prélèvements pharyngés ou de selles de patients hospitalisés entre 2000 et 2013 et se répartissent dans 9 sérotypes (Tableau 6). Toutefois, les souches de la Fédération de Russie, d'Azerbaïdjan, de Kazakhstan et de Kirghizistan n'ont pas été génotypées au laboratoire.

### **2) Echantillon constitué pour étudier l'évolution de la capsid de l'EV-A71**

L'évolution moléculaire de la capsid de l'EV-A71 a été étudiée avec un échantillon constitué de 126 souches isolées à partir de prélèvements pharyngés ou de selles de patients hospitalisés entre 1987 et 2014 dans quatre pays européens : France (n=84), Allemagne (n=33), Hongrie (n=6) et Autriche (n=3).

### **3) Echantillon constitué pour étudier la recombinaison chez les EV-A**

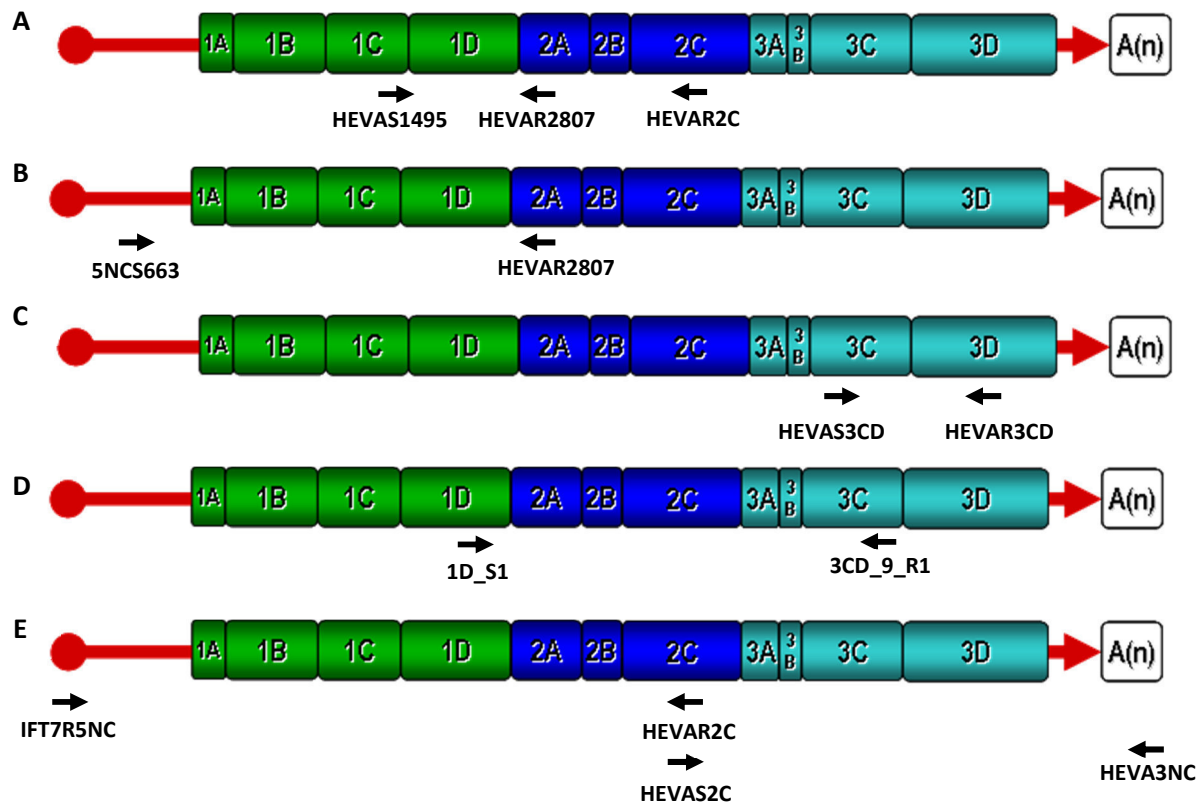
Cette étude a été effectuée avec 97 souches d'EV-A sélectionnées parmi les deux échantillons décrits ci-dessus : France (n=73), Allemagne (n=15), Hongrie (n=4), Autriche (n=3), Finlande (n=1) et Islande (n=1).

### **4) Echantillon constitué pour étudier le pouvoir infectieux des EV-A**

Quarante et une souches ont été sélectionnées parmi l'échantillon décrit ci-dessus afin de représenter la diversité génétique des EV-A. Elles proviennent de prélèvements pharyngés ou de selles d'enfants atteints d'un syndrome PMB en France (n=29), Allemagne (n=7), Finlande (n=2), Hongrie (n=2) et Autriche (n=1).

### **5) Isolement des souches et titrage du pouvoir infectieux**

La lignée cellulaire RD (rhabdomyosarcome humain) a été utilisée pour isoler les souches d'entérovirus et préparer les stocks nécessaires à cette étude. Les cellules ont été cultivées entre les passages 45 et 70 dans un milieu de culture complet : milieu RPMI 1640 (Eurobio)



**Figure 36. Structure schématique du génome d'un entérovirus et la localisation des amorces employées dans les différentes réactions d'amplification génique**

A : amplification du locus 1D<sup>VP1</sup> (amplicons de 1200 et 2200 pb)

B : amplification de la région génome P1 (amplicons de 2900 pb)

C : amplification du locus 3CD (amplicons de 1200 pb)

D : amplification du subgénome entre les locus 1D<sup>VP1</sup> et 3CD (amplicons de 3000 pb)

E : amplification du génome complet en deux parties (amplicons de 4400 et 3000 pb)

supplémenté avec 1% de pénicilline-streptomycine (PAA) et de 4 % de sérum de veau fœtal (Eurobio).

Le titre infectieux (nombre de particules virales infectieuses par unité de volume) a été déterminé par une micro méthode en plaques de culture à 96 puits mise au point et utilisée en routine dans le laboratoire basée la méthode des dilutions limites. Les plaques sontensemencées avec les cellules RD à 80000 cellules par mL, cultivées dans le milieu RPMI 1640 complet. L'inoculation des cellules est effectuée 48 heures après avec différentes dilutions de la suspension virale à titrer. L'inoculum a été préparé selon la méthode décrite dans les travaux de Bailly *et al*, 1991.

#### **6) Désignation des souches**

Les souches ont été désignées de la façon suivante : l'abréviation du sérotype, les premières lettres indiquent la ville, les chiffres suivant le numéro du prélèvement dans le laboratoire qui a effectué l'isolement, puis trois lettres pour indiquer le pays d'origine (code ISO alpha 3) et enfin l'année d'isolement. Par exemple, la désignation CVA4\_CF1188031\_FRA12 correspond à la souche de coxsackievirus A4 n°188031 isolée Clermont-Ferrand, France, en 2012.

## **II. Extraction de l'ARN viral et synthèse d'ADNc**

Les ARN génomiques viraux contenus dans 200 µL de surnageant de culture ont été extraits en utilisant une méthode semi automatisée avec l'automate NucliSENS® easyMAG™ (bioMérieux, France). Les acides nucléiques totaux contenant l'ARN viral ont été élués dans 25 µL du tampon d'éluion et la solution conservée à -80°C.

La synthèse d'ADNc a été réalisée en utilisant des hexanucléotides comme amorces et l'enzyme Superscript III (Invitrogen, Cergy-Pontoise, France). L'hybridation des amorces a été effectuée pendant 5 min à 65°C et l'élongation pendant 50 min à 50°C (Mirand *et al*, 2008).

## **III. Amplification par RT-PCR de différents fragments génomiques**

### **1) Séquence génique 1D<sup>VP1</sup>**

Toutes les souches étudiées ont été identifiées par géotypage en amplifiant la séquence du



**Tableau 7. Conditions d'amplification génique selon les amorces utilisées**

Amorce	Spécificité	Région	Tm (°C)	Séquence (5' → 3') <sup>1</sup>
IFT7R5NC	EV	5NC	62.90	TTAAAACAGCCTGTGGGTTG
5NCS663	EV	5NC	76.69	GCGGAACCGACTACTTTGGGTGTCCGTGTTTC
HEVAS1495	EV-A	1C	58.39	GGTNATNCCNTGGATAAGYACAC
CA2VP1_S1	CV-A2	1D	57.30	GCTAGTGCATACCAGTGG
CA4-VP1-S1	CV-A4	1D	70.80	CCRCCACCACARGTGTGTCAG
CA6VP1_S1	CV-A6	1D	56.80	GTGTAYATGAGAATTAARCACG
CA8-VP1_S1	CV-A8	1D	61.30	GCCTATTTCGATCACAGCC
CA10_1D_S1	CV-A10	1D	75.80	ATGATGGGYACYTTTGCAGTGCG
CA16_1D_S1	CV-A16	1D	70.30	TGYCCGAAYAAYATGATGGGCAC
1D-6-S1	EV-A71	1D	72.40	RCACGTCAGGGCGTGGRTAC
HEVAR2807	EV-A	2A	59.09	CYACNGCNCARGGNTGYGACACGA
HEVAR2C	EV-A	2C	68.07	CATGCAGTTCAAGAGCAARCACCG
HEVAS2C	EV-A	2C	68.07	CATGCAGTTCAAGAGCAARCACCG
3CD-9-R1	EV-A	3C	56.98	GAGCACATGCCMTCAATG
HEVAS3CD	EV-A	3C	66.62	GACCARGGNCAYTTYACNATGYTRGG
HEVAR3CD	EV-A	3D	48.60	GGAGYAARYTACCRATYCTA
HEVA3NC	EV-A	3NC	68.70	GTGGGGGTAAATTTGTTATAACCAGAATAGC

<sup>1</sup> N = inosine, R = A ou G, Y = C ou T, M = A ou C

**Tableau 8. Conditions d'amplification des séquences géniques 1D, P1 et 3CD avec l'enzyme polymérase Platinum Taq (Invitrogen)**

Région amplifiée	1D	1D	P1	3CD
Amorce 1	HEVAS1495	HEVAS1495	5NCS663	HEVAS3CD
Amorce 2	HEVAR2C	HEVAR2807	HEVAR2807	HEVAR3CD
Dénaturation initiale	94°C ; 2 min	94°C ; 2 min	94°C ; 2 min	94°C ; 2 min
Dénaturation	94°C ; 15 sec	94°C ; 15 sec	94°C ; 15 sec	94°C ; 15 sec
Hybridation	53°C ; 50 sec	50°C ; 50 sec	50°C ; 50 sec	53°C ; 50 sec
Elongation	68,5°C ; 90 sec	72°C ; 50 sec	72°C ; 105 sec	68,5°C ; 90 sec
Nombre de répétitions	41	39	39	41
Elongation finale	68,5°C ; 5 min	72°C ; 5 min	72°C ; 5 min	68,5°C ; 5 min
Taille de l'amplicon	2219 pb	1273 pb	2650 pb	1960 pb

**Tableau 9. Conditions d'amplification du génome complet en deux étapes avec la polymérase Phusion Flash (Thermoscientific)**

Conditions d'amplification	Partie 5' du génome	Partie 3' du génome
Amorce 1	IFT7R5NC	HEVAS2C
Amorce 2	HEVAR2C	HEVA3NC
Dénaturation initiale	98°C ; 10 sec	98°C ; 10 sec
Dénaturation	98°C ; 5 sec	98°C ; 5 sec
Hybridation et élongation	72°C ; 70 sec	72°C ; 70 sec
Nombre de répétitions	41	41
Elongation finale	72°C ; 2 min	72°C ; 2 min
Taille de l'amplicon	4450 pb	3000 pb

locus 1D<sup>VP1</sup> par PCR simple ou semi-nichée (**Figure 36A**). L'amplification génique simple a été réalisée avec les amorces HEVA1495 et HEVAR2807 ; lorsque la sensibilité était insuffisante, une amplification semi nichée préalable a été réalisée avec les amorces HEVAS1495 et HEVAR2C (**Tableau 7**). L'enzyme Platinum Taq Polymerase (Invitrogen) a été utilisée en suivant les conditions décrites dans le **Tableau 8**.

## 2) Région génomique P1

Les réactions pour l'amplification génique de la région P1 ont été réalisées avec les amorces 5NCS663 et HEVAR2807 (**Figure 36B**) qui permettent d'obtenir un amplicon d'environ 2650 nucléotides (**Tableaux 7 et 8**).

## 3) Séquence génique partielle 3CD

Les réactions d'amplification génique d'un fragment du locus 3CD ont été effectuées avec les amorces HEVAS3CD et HEVAR3CD (**Figure 36C ; Tableaux 7 et 8**).

## 4) Fragment subgénomique compris entre 1D<sup>VP1</sup> et 3CD

L'amplification des fragments subgénomiques incluant les séquences génomiques comprises entre les locus 1D<sup>VP1</sup> et 3CD (**Figure 36D**) a été réalisée avec la polymérase Phusion Flash (Thermoscientific). Les couples d'amorces utilisées varient selon les sérotypes (**Tableau 7**) ; les conditions d'amplification sont décrites dans le **tableau 8**.

## 5) Génome complet

L'amplification du génome complet des souches d'EV-A a été réalisée avec la polymérase Phusion Flash (Thermoscientific) en deux étapes (**Figure 36E**). La première a consisté à amplifier un fragment de la partie 5' du génome avec les amorces IFT7R5NC et HEVAR2C ; la seconde étape couvre le reste du génome entre le gène 2C et la région 3'UTR en utilisant les amorces HEVAS2C et HEVA3NC. Les caractéristiques des amorces et les conditions d'amplifications employées sont décrites dans les **Tableaux 7 et 9**.

# IV. Purification et séquençage des produits d'amplification

Les produits de PCR ont été contrôlés par électrophorèse en gel d'agarose (1%) avec le marqueur fluorescent Gelstar, avant d'être purifiés (NucleoSpin gel extraction Kit, Macherey-Nagel). La solution d'ADN purifié a été conservée à  $-20^{\circ}\text{C}$ . Les réactions de séquence ont été

**Tableau 10. Constitution des jeux de données de séquences du locus 1D<sup>VP1</sup> utilisés pour les analyses phylogénétiques**

Sérotype	Genogroupe	Total <sup>a</sup>	Laboratoire <sup>b</sup>	GenBank <sup>c</sup>
EV-A		315	77	238 (17)
CV-A2		48	18	31
CV-A4		75	33	42
CV-A5		17	9	8
CV-A6		617	126	491
CV-A8		25	13	12
CV-A10		201	38	163 (40)
CV-A14		18	8	10
CV-A16		572	157	415
EV-A71	<b>B</b>	438	1	436 (1)
EV-A71	<b>B4/B5</b>	248	1	247
EV-A71	<b>C1</b>	280	1	279 (42)
EV-A71	<b>C2</b>	322	55	267 (35)
EV-A71	<b>C4</b>	675	8	667 (6)

<sup>a</sup> Nombre total de séquences nucléotidiques analysées dans cette étude.

<sup>b</sup> Nombre de séquences nucléotidiques déterminées au cours de l'étude.

<sup>c</sup> Nombre des séquences nucléotidiques collectées dans la base de données GenBank (NCBI). Le nombre de séquences déposées antérieurement par le laboratoire est indiqué entre parenthèses.

effectuées avec le kit BigDye terminator v1.1 cycle sequencing (Applied Biosystems, Foster City, USA) et les amorces de l'amplification génique (Mirand *et al*, 2010). Les amorces intermédiaires (EV71-C1-564, EV71-C1-1160 et HEVR436 ont été utilisées pour déterminer la séquence de la région P1. D'autres amorces intermédiaires (non décrites) ont été utilisées pour séquencer les subgénomés et les génomes complets. Les produits des réactions de séquence ont été analysés par électrophorèse capillaire (ABI 3500 DX ; Applied Biosystems, Foster City, USA). Les chromatogrammes ont été traités avec le logiciel BioEdit v7.0.9 (Hall, 1999) pour obtenir les séquences « contigs ».

## **V. Constitution des jeux de séquences nucléotidiques (Tableau 10)**

### **1) Analyses préalables des jeux de séquences**

Les séquences nucléotidiques ont été alignées avec l'algorithme CLUSTALW 2.0 (Larkin *et al*, 2007) implémenté dans le programme BioEdit v7.0.9. Les jeux de données destinés à des analyses bayésiennes ont été soumis à une analyse avec les méthodes des logiciel RDP4 (Martin *et al*, 2010) et SBP disponible sur la plateforme web Datamonkey (Kosakovsky Pond *et al*, 2006) afin de détecter d'éventuels événements de recombinaison. Les séquences qui présentaient des sites de recombinaison ont été exclues.

### **2) Jeux des séquences 1D<sup>VP1</sup>**

Les séquences nucléotidiques 1D<sup>VP1</sup> déterminées ont été assemblées à celles collectées dans la base de données GenBank (dernière consultation : août 2014) afin d'accroître l'échantillonnage temporel et géographique des jeux de données. Concernant l'EV-A71, seules les séquences nucléotidiques complètes ont été utilisées et des jeux de séquences ont été constitués pour analyser séparément le génogroupe B et les sous-génogroupes B4-B5, C1, C2 et C4.

### **3) Jeu des séquences P1 constitué pour l'EV-A71**

Les séquences nucléotidiques P1 déterminées chez les souches d'EV-A71 étudiées ont été assemblées à celles collectées dans GenBank (dernière consultation, 29 août 2014). Les séquences incomplètes (taille < 2586 nucléotides) et les séquences complètes qui ne disposaient pas d'information concernant l'année d'isolement n'ont pas été utilisées. Le jeu de données final contenait 442 séquences.

#### **4) Jeux des séquences 3CD**

Les séquences nucléotidiques 3CD déterminées pour cette étude ont été assemblées à celles de GenBank (dernière consultation, 16 juin 2013). La taille minimum des séquences 3CD était de 650 nucléotides.

#### **5) Jeux des séquences subgénomiques et des génomes complets**

Les séquences de génomes complets et des subgénomiques déterminées pour cette étude ont été assemblées à celles collectées dans Genbank (dernière consultation, 5 juin 2014). Les parties codantes des séquences alignées (polyprotéine) ont été converties en acides aminés et réalignées en utilisant l'algorithme MUSCLE (implémenté dans le programme Seaview v4.0 ; <http://doua.prabi.fr/software/seaview>). Puis, l'alignement obtenu a fait l'objet d'une transcription inverse en nucléotides.

### **VI. Analyse de la diversité des EV-A**

La diversité des lignées a été analysée pour les jeux de données CV-A2, CV-A4, CV-A5, CV-A6, CV-A8, CV-A10, CV-A14 et CV-A16. La proportion de différences nucléotidiques (modèle P-distance) a été estimée et représentée sous la forme d'un arbre en utilisant le logiciel MEGA 5 (Tamura *et al*, 2011). La fiabilité de la topologie phylogénétique a été vérifiée par la méthode du « bootstrap » (n = 1000 répétitions). Une valeur de bootstrap > 70% caractérise un nœud fiable.

### **VII. Analyse exploratoire de la recombinaison**

#### **1) Recherche de souches d'origine recombinante par différences de topologies**

Une analyse exploratoire a été réalisée afin d'identifier des souches d'origine recombinante. Ce criblage consiste à inférer les phylogénies de deux gènes distants (loci 1D<sup>VP1</sup> et 3CD) pour le même échantillon de séquences. Les incongruences des topologies suggèrent des événements de recombinaison. Les arbres phylogénétiques ont été construits avec le logiciel phyML par la méthode du maximum de vraisemblance avec le modèle d'évolution moléculaire GTR. La fiabilité des embranchements a été évaluée par la méthode du Bootstrap (1000 répétitions).

### **2) Analyse de la similarité nucléotidique des subgénomés / génomes complets**

Le logiciel Simplot v3.5.1 (Lole *et al*, 1999) a été utilisé pour déterminer les profils de similitudes nucléotidiques entre les génomes complets et les séquences subgénomiques d'EV-A (n=226) et d'EV-A71 (n=341). La similitude a été calculée avec le modèle de distance génétique Kimura à 2 paramètres avec une fenêtre d'analyse de 200 nucléotides glissant par pas de 20 nt le long de l'alignement de séquences multiples. Les sites de recombinaison ont été localisés à l'aide de la fonction Bootscan. Les arbres phylogénétiques ont été estimés avec la méthode « Neighbour Joining » et un Bootstrap de 200 répétitions.

### **3) Analyse des signatures moléculaires de recombinaison**

Le logiciel RDP3 a été utilisé pour détecter des signatures moléculaires caractéristiques de la recombinaison, les sites de recombinaison (Martin, 2009). Les algorithmes de 7 méthodes d'analyse sont implémentés : RDP (Martin and Rybicki, 2000), GENECONV (Padidam *et al*, 1999), Bootscan (Martin *et al*, 2005), MaxChi (Smith, 1992), Chimera (Posada et Crandal, 2001), Siscan (Gibbs *et al*, 2000) et 3SEQ (Boni *et al*, 2007). Les options par défaut ont été employées pour les analyses. Les sites de recombinaison estimés par moins de 3 méthodes n'ont pas été considérés.

## **VIII. Analyse du signal temporel dans les jeux de séquences**

Avant l'investigation par une approche bayésienne, un arbre phylogénétique a été reconstruit pour chaque jeu de séquences avec la méthode du maximum de vraisemblance. Un modèle de régression linéaire a été appliqué pour analyser la relation entre les distances génétiques estimées entre les feuilles et la racine de l'arbre et l'année d'isolement des souches. L'analyse a été effectuée avec le programme Path-O-Gen (<http://tree.bio.ed.ac.uk/software/pathogen/>).

## **IX. Analyse phylogénétique par une méthode bayésienne**

Les généalogies des différentes populations virales étudiées ont été reconstruites par une méthode bayésienne en utilisant le logiciel BEAST v1.7.5 (Drummond *et al*, 2012). Les généalogies et les densités de probabilité postérieure des valeurs numériques des paramètres phylogénétiques et démographiques ont été estimées en utilisant 21 combinaisons de modèles

d'horloge moléculaire, d'évolution moléculaire et de démographie. Ces analyses ont été réalisées avec des chaînes de Markov de Monte Carlo (MCMC) dont la longueur variait entre 100 millions et 200 millions de générations pour obtenir des estimations statistiquement fiables. Les chaînes MCMC étaient programmées de façon à obtenir un échantillon final de 20 000 arbres phylogénétiques. La progression des chaînes de Markov et la convergence des paramètres phylogénétiques ont été suivies grâce au logiciel Tracer v1.5.0 (Drummond & Rambaut, 2007). Implémenté dans Tracer, le test statistique « Bayes Factor » a été utilisé pour comparer les combinaisons de modèles et déterminer celle dont la distribution de probabilité postérieure représentait le mieux les données. La généalogie dont la fonction de vraisemblance était la plus élevée « maximum clade credibility (MCC) tree » a été reconstruite avec le logiciel TreeAnnotator v1.7.5 (Drummond & Rambaut, 2007). Deux arbres phylogénétiques ont été examinés pour chaque jeu de données : le phylogramme dans lequel la longueur des branches est proportionnelle aux distances génétiques et le chronogramme dans lequel l'âge de l'ancêtre commun le plus récent « time to the most recent common ancestor » (TMRCA) et celui de chaque nœud est estimé. Les arbres ont été visualisés avec le logiciel FigTree v1.4.0 (Drummond & Rambaut, 2007).

## **X. Reconstruction de l'histoire démographique des populations virales**

La diversité génétique relative des populations virales au cours du temps a été estimée avec le modèle Bayesian Skyline Plot (BSP). La diversité génétique est le produit de la taille efficace de la population ( $N_e$ ) et le temps de génération ( $t$ ), représentant le nombre effectif d'infections moyennes au cours du temps sous la contrainte d'un processus évolutif neutre. L'analyse du BSP a été effectuée avec un modèle d'horloge moléculaire assouplie dans lequel les taux d'évolution évoluent selon une loi log-normale et la taille des populations au cours du temps était interpolée directement sur la phylogénie selon la méthode « piecewise-constant ».

## **XI. Estimation de l'origine géographique des nœuds ancestraux**

Les pays dans lesquels les souches virales ont été isolées ont été employés comme état de caractère discret pour estimer la probabilité de l'origine géographique des nœuds dans les

arbres phylogénétiques inférés (chronogrammes). Un modèle discret de diffusion géographique « symmetrical substitution model » a été utilisé pour cette analyse. L'algorithme « Bayesian Stochastic Search Variable Selection ; BSSVS » a permis d'identifier les nœuds qui présentaient des changements d'origine géographique significatifs.

## **XII. Polymorphisme des protéines de capsid**

### **1) Calcul de l'entropie de Shannon**

L'entropie de Shannon est une mesure quantitative de la variabilité de chaque position d'un alignement de séquences nucléiques ou protéiques. L'entropie de Shannon (Schmitt Herzel, 1997) a été calculée pour mesurer la variabilité des quatre protéines de capsid de l'EV-A71. L'algorithme de la méthode "ENTROPY-TWO", est implémenté sur la plateforme "HIV sequence database" (<http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html>) et permet de comparer la variation protéique entre deux alignements différents.

### **2) Reconstruction des états ancestraux**

Pour caractériser l'évolution moléculaire qui a généré les différentes populations virales de l'EV-A71, les mutations survenues dans les lignées phylogénétiques ont été déterminées en utilisant une généalogie estimée par l'approche bayésienne. Les séquences nucléotidiques de chaque nœud ancestral de l'arbre établi pour la région génomique P1 ont été reconstruites avec la méthode probabiliste « Ancestral State Reconstruction » (ASR). Cette méthode permet d'inférer les séquences nucléotidiques les plus probables d'un nœud en calculant la probabilité des nucléotides à chaque position d'une séquence. Il s'agit d'une méthode de reconstruction « joint » qui tient compte du taux de variations des séquences protéiques selon une distribution gamma le long des positions de l'alignement. A partir de l'ensemble des probabilités à un nœud précis, il est possible d'en inférer la séquence ancestrale. Un algorithme de cette méthode est disponible sur la plateforme web « FastML computing Maximum Likelihood ancestral sequence reconstruction » (Pupko, 2000 ; Ashkenazy, 2010).



### **XIII. Estimation des pressions de sélection dans les protéines de capsid**

La sélection naturelle qui opère sur les souches joue un rôle majeur dans l'évolution des populations virales. Les différentes formes de sélection peuvent laisser des signatures moléculaires à chaque codon. Les méthodes disponibles sur la plateforme web HYPHY/Datamonkey ont été utilisées pour cette analyse (Delpont, 2010). Le modèle de substitutions nucléotidiques le plus adapté au jeu de données a été déterminé et utilisé pour l'ensemble des analyses sur cette plateforme. Le rapport des substitutions non synonymes et des substitutions synonymes ( $dN/dS$  ou  $\omega$ ) et les sites sélectionnés négativement ou positivement ont été estimés par six méthodes (Kosakovsky Pond, 2005). La méthode « single-likelihood ancestor counting » (SLAC) permet de calculer le nombre global de substitutions synonymes et non synonymes à chaque site en utilisant les états ancestraux de la phylogénie du jeu de données. La méthode « fixed effects likelihood » (FEL) estime ces deux taux de la même manière, site par site, en considérant que les taux sont indépendants. La méthode « internal fixed effects likelihood » IFEL estime «  $\omega$  » de la même manière que la méthode FEL uniquement sur les branches internes de l'arbre phylogénétique. La méthode bayésienne hiérarchique « fast unconstrained Bayesian approximation » (FUBAR) permet d'estimer une distribution postérieure du taux «  $\omega$  » à un site donné. La méthode MEME « Mixed Effects Model of Evolution » fait varier le taux «  $\omega$  » le long des branches de l'arbre phylogénétique suivant une distribution à 2 données permettant la détection des sites soumis à une sélection épisodique. Enfin, l'analyse PRIME « PRoperty Informed Models of Evolution » utilise la méthode que le FEL en considérant les propriétés biochimiques des résidus substitués.

### **XIV. Numéros d'accèsion des séquences nucléotidiques**

Une partie des séquences nucléotidiques déterminées dans cette étude a été déposée dans les bases de données GenBank/EMBL/DDBJ sous les numéros d'accèsion HG934162-HG934296. Les autres séquences nucléotidiques déterminées au cours de ces travaux seront déposées ultérieurement sur Genbank.

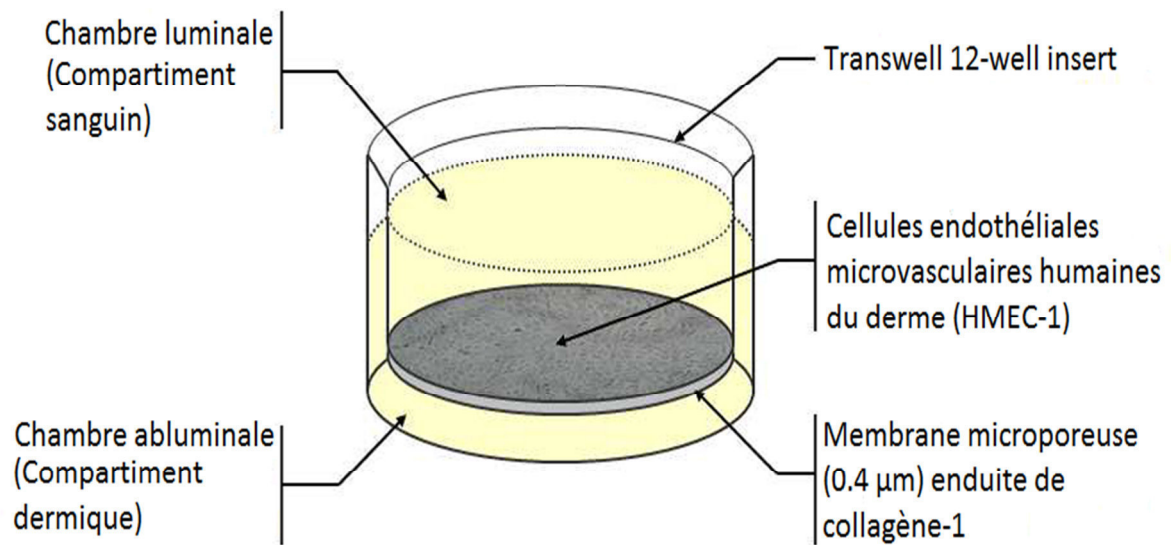
## **XV. Etude de l'interaction des entérovirus avec un modèle de barrière endothéliale de derme**

### **1) Lignée HMEC-1 (Human Microvascular Endothelial Cell)**

La lignée de cellules endothéliales microvasculaires de derme humain (Ades *et al*, 1992) a été obtenue auprès du Centre de Contrôle des Maladies Infectieuses (CDC, Atlanta, USA) sous un contrat de transfert de technologies. Cette lignée a été cultivée sur un support recouvert de collagène et avec un milieu complet spécifique des cellules endothéliales MV2 (Promocell, Allemagne) et supplémenté avec 1% de pénicilline-streptomycine. Les cellules ont été utilisées entre passages 12 et 25 pour les expérimentations.

### **2) Sensibilité de la lignée HMEC-1 à l'infection par les entérovirus de l'espèce A**

Les cellules HMEC-1 ( $4 \times 10^4$ ) ont été cultivées à la surface de lames en verre compartimentées en 8 puits, dont la surface était pré-induite de collagène I. Elles ont été incubées pendant 48h à 37°C (atmosphère à 5% CO<sub>2</sub>) pour atteindre la confluence. Les monocouches cellulaires ont été lavées 2 fois avec du PBS et inoculées avec une souche virale au cours d'une incubation d'une heure à 37°C pour initier l'infection. L'inoculum viral a été prélevé et les cellules lavées deux fois avec du PBS. Les tapis cellulaires infectés ont été incubés durant 48h. La formation d'un ECP a été observée chaque jour au microscope optique. Les tapis cellulaires infectés ont été fixés avec une solution de para formaldéhyde à 4% pendant cinq minutes à température ambiante puis lavés avec une solution de PBS. Les cellules ont été perméabilisées au cours d'une incubation de 10 minutes à température ambiante avec une solution de Triton X100 à 0.5%. Les sites de fixation non spécifique des anticorps ont été saturés avec une solution de BSA à 5% pendant 10 minutes à température ambiante. Les tapis cellulaires ont été incubés (une nuit à 4°C) avec la solution de l'anticorps primaire : l'anticorps monoclonal dirigé contre la protéine de capsid VP1 (Diagnostic Hybrid) ou l'anticorps dirigé contre les ARN double brin viraux (English Scientific Consulting). Après trois lavages avec du PBS, les cellules ont été incubées pendant une heure à 37°C avec une solution de l'anticorps secondaire « anti-souris » conjugué au fluorophore Dyelight 488 (Anticorpsenligne). L'ADN nucléaire des cellules a été marqué avec une solution de Hoechst durant 5 minutes à température ambiante. Les tapis cellulaires ont été recouverts d'une solution de glycérine/PBS (v/v) et d'une lamelle de protection. Les observations ont été effectuées avec un microscope à épifluorescence Olympus BX41 et avec



**Figure 37. Schéma du modèle *in vitro* de barrière endothéliale microvasculaire de derme**

Les cellules de lignées HMEC-1 sont cultivées au sein d'un insert de culture cellulaire, modèle Transwell 12 puits. Les cellules sont présentes à la surface d'une membrane microporeuse (pores de 0,4 µm de diamètre), enduite de collagène I. Selon ce modèle, la chambre luminale mime le compartiment sanguin et la chambre abluminale représente le compartiment dermique.

un microscope confocal Leica SPE manual, disposant de lasers Argon (488 nm), HeNe568 (568 nm) et HeNe633 (633 nm).

### 3) Modèle *in vitro* de barrière endothéliale microvasculaire de derme

Les cellules HMEC-1 ( $4 \times 10^4$ ) ont été cultivées sur une membrane microporeuse (pores de 0.4  $\mu\text{m}$  de diamètre), installée dans un dispositif (nacelle) Transwell® 12 puits (Corning) et pré-induite de collagène I. Une monocouche cellulaire confluite est obtenue après une incubation de 72h à 37°C dans une atmosphère contenant 5% de CO<sub>2</sub> (**Figure 37**).

L'étanchéité paracellulaire de la monocouche a été estimée en déterminant le coefficient de perméabilité d'un fluorophore, le Lucifer Yellow (LY, Sigma Aldrich). Deux lavages des compartiments luminaux et abluminaux ont été effectués avec du tampon HBSS (1% d'HEPES à 10 mM et 1% de pyruvate de sodium à 1 mM). La solution du fluorophore LY (50  $\mu\text{M}$ ) a été déposée dans le compartiment luminal et la nacelle a été incubée à 37 °C successivement dans trois puits indépendants contenant du tampon HBSS, pendant 10 min (puits 1), 25 min (puits 2) et 45 min (puits 3). La quantité de LY qui a traversé la monocouche cellulaire HMEC-1 par voie paracellulaire a été déterminée dans l'éluat recueilli dans les trois puits. La quantité résiduelle de LY dans le compartiment luminal (temps 45 minutes) est également déterminée. Le dosage du LY a été effectué en plaque 96 puits opaque en mesurant la fluorescence émise (excitation 425 nm, émission à 538 nm) à l'aide d'un fluorimètre (Fluoroskan, Thermo Scientific). Le coefficient de perméabilité paracellulaire du LY a été calculé par unité de temps et de surface.



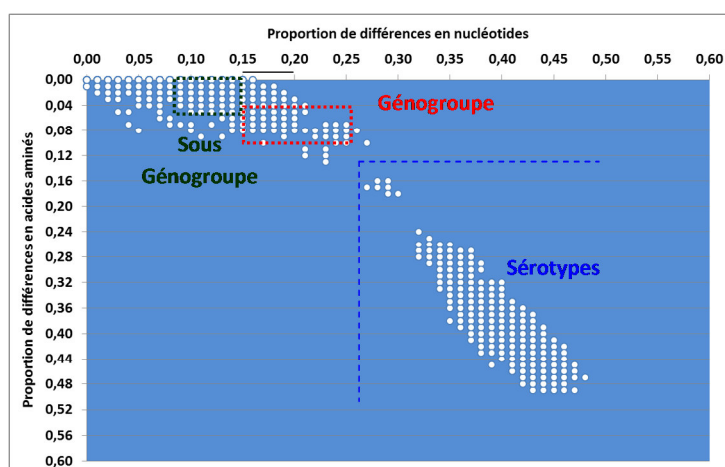
**TRAVAUX**

**PERSONNELS**

# I. Epidémiologie moléculaire comparée des sérotypes d'EV-A

## 1) Polymorphisme des séquences moléculaires et diversité intratypique

Dans une même espèce taxonomique, les niveaux de classification des souches d'entérovirus employés couramment sont le sérotype, le génogroupe (ou génotype) et le sous-génogroupe (ou sous-génotype). Pour comparer la diversité des souches d'EV-A, la distribution des différences nucléotidiques et protéiques a été analysée graphiquement en utilisant un échantillon de 315 séquences [Figure 38]. La distribution se caractérise par deux nuages de points principaux suggérant l'existence de populations dans l'échantillon. L'identification des points dans chaque nuage montre qu'ils correspondent à des comparaisons de séquences représentant des niveaux taxonomiques différents (sérotype, génogroupe, sous-génogroupe). Des valeurs seuils de différences nucléotidiques et protéiques peuvent être définies : deux souches de l'espèce EV-A appartiennent à deux sérotypes différents dès lors qu'elles présentent entre 25 % et 50 % de différences nucléotidiques et entre 13 % et 50 % de différences protéiques [Tableau 11].



**Figure 38. Analyse graphique des différences nucléotidiques en fonction des différences protéiques du gène 1D<sup>VP1</sup> parmi les souches d'EV-A**

Les séquences du gène 1D<sup>VP1</sup> (jeu de données HEVA/1D/233) ont été comparées pour un échantillon de 315 souches d'EV-A. Une matrice des proportions de différences nucléotidiques (p-distance) et une autre matrice des proportions de différences en acides ont été obtenues avec le logiciel MEGA5 (Tamura *et al*, 2007) pour comparer les souches deux à deux. La distribution des valeurs a été analysée graphiquement afin d'identifier l'existence de groupe (populations) parmi les souches et d'établir les seuils de différenciation entre sérotypes, génogroupes et sous génogroupes.

**Tableau 11. Valeurs seuils pour différencier les niveaux de classifications taxonomiques**

	Pourcentage de différences nucléotidiques	Pourcentage de différences protéiques
<b>Sous génogroupes</b>	<b>10 - 14 %</b>	<b>0 - 6 %</b>
<b>Génogroupes</b>	<b>15 - 25 %</b>	<b>2 - 10 %</b>
<b>Sérotypes</b>	<b>26 - 50 %</b>	<b>13 - 50 %</b>





Avec ces valeurs seuils, la diversité intratypique a été analysée en utilisant les jeux de données 1D<sup>VP1</sup> constitués pour les sérotypes CV-A2, CV-4, CV-5, CV-6, CV-A8, CV-A10, CV-A14 et CV-A16 et les génogroupes et de sous-génogroupes ont été définis pour chacun. La diversité génétique la plus importante (mesurée par le nombre de génogroupes et de sous-génogroupes) est observée parmi les sérotypes CV-A2, CV-A4, CV-A6, CV-A10 et CV-A14 [Figure 39]. Avec deux génogroupes, le sérotype CV-A16 constitue une exception puisqu'il est relié à une intense circulation des souches (nombreuses épidémies de maladie pied-main-bouche en Asie). Le sérotype CV-A14 constitue une exception inverse : il comprend 5 génogroupes alors que ce sérotype est rapporté dans très peu d'infections symptomatiques. Il est vraisemblable que ces exceptions ne soient qu'apparentes, en raison du grand nombre des infections asymptomatiques et pauci-symptomatiques, non rapportées.

## **2) Epidémiologie moléculaire de l'EV-A71**

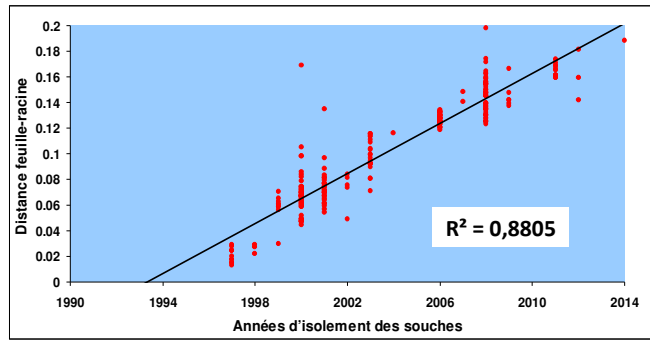
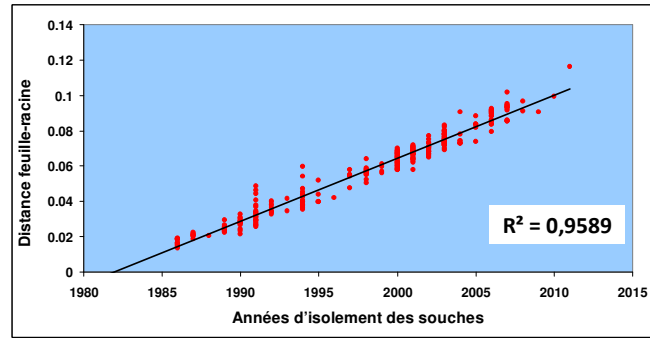
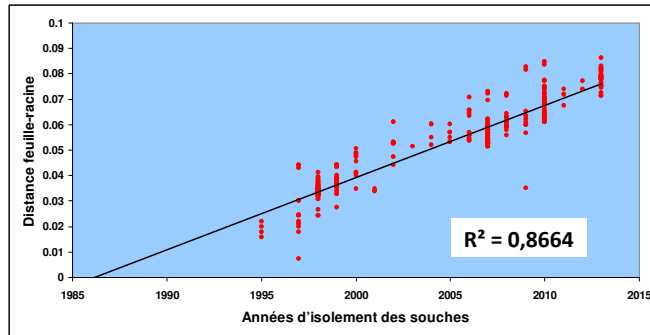
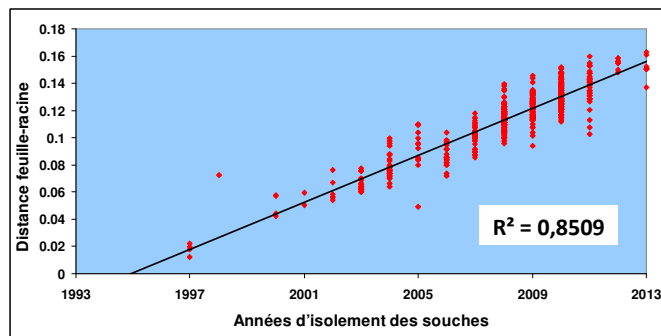
Les données décrites dans cette partie ont fait l'objet de deux articles, (i) un centré sur les profils de transmission de l'EV-A71 entre les différents pays européens entre 2003 et 2013 dans le journal « Eurosurveillance », accepté le 21 janvier 2015 et (ii) un autre décrivant l'histoire évolutive du premier variant du génogroupe B5 d'EV-A71 mise en évidence en France sous la forme d'une lettre dans le journal « Emerging Infectious Diseases », accepté le 9 janvier 2015.

### **A. Evaluation de l'horloge moléculaire dans les jeux de données de l'EV-A71**

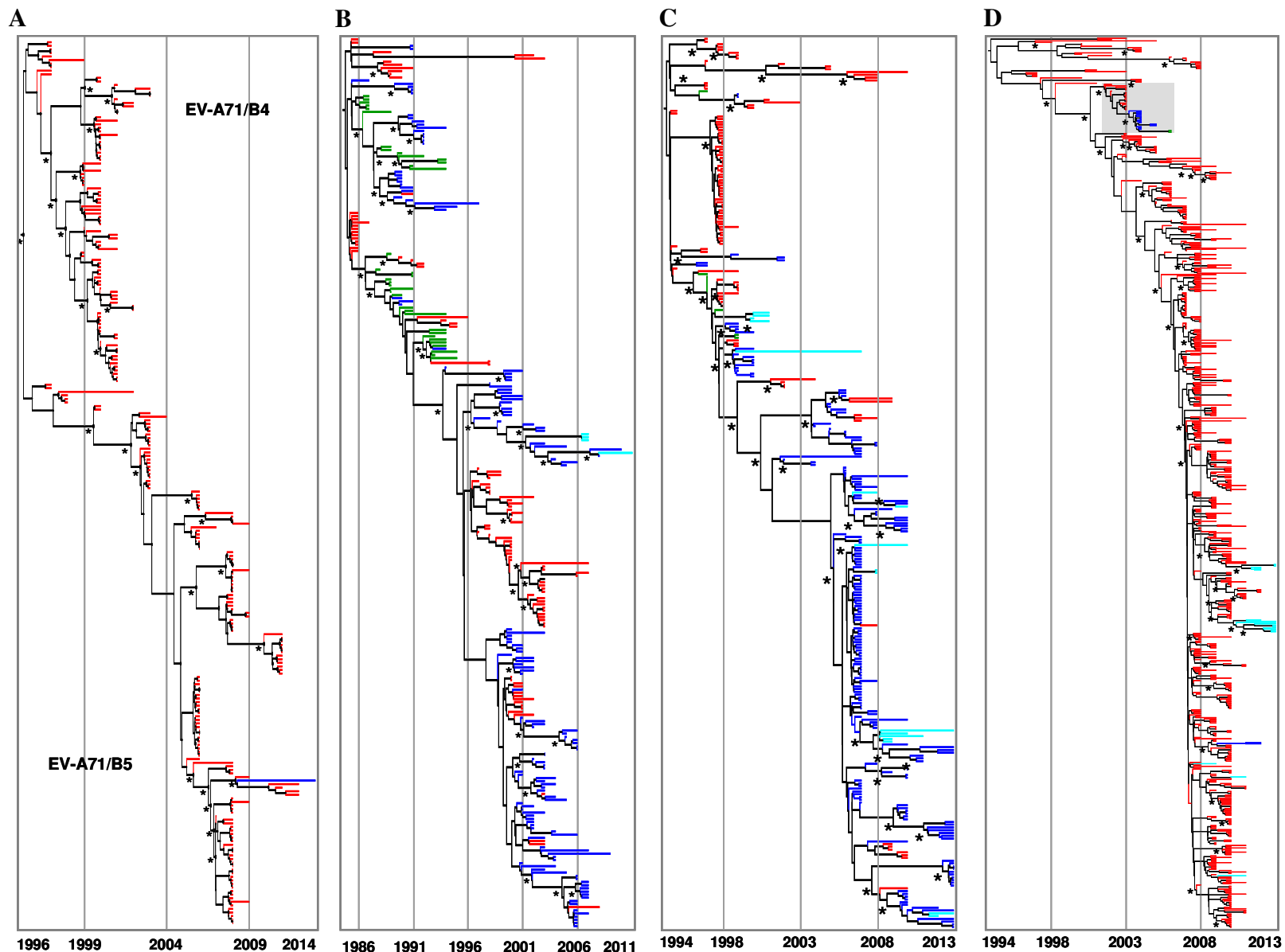
Un signal temporel (horloge moléculaire) a été recherché dans les jeux de données constitués pour comparer les sous-génogroupes d'EV-A71 en utilisant les arbres phylogénétiques inférés par maximum de vraisemblance. Une analyse de régression linéaire des distances « feuille / racine » a été effectuée [Figure 40]. Les résultats montrent l'existence d'une relation entre les distances génétiques et les dates d'isolement des souches pour tous les jeux de données analysées avec des coefficients de détermination  $R^2$  compris entre 0,85 (EV71-C4) et 0,96 (EV71-C1). Un modèle d'horloge moléculaire assoupli peut donc être utilisé pour analyser les jeux de données avec une approche bayésienne.

### **B. Profils phylogénétiques des populations d'EV-A71 reliées à l'épidémiologie**

Dans un premier temps, le modèle d'analyse le plus adéquat à chaque jeu de données a été recherché en comparant 21 combinaisons de modèles d'évolution moléculaire, d'horloge

**A****B****C****D**

**Figure 40. Régression linéaire de la distance feuille-racine par rapport à l'année d'isolement des souches.** Les analyses **a**, **b**, **c**, **d** ont été effectuées respectivement avec les arbres phylogénétiques des jeux de données EV-A71/B4-B5 ; EV-A71/C1 ; EV-A71/C2 et EV-A71/C4. Une relation linéaire entre les distances génétiques (distances racine-feuille d'un arbre non raciné) et le temps (année d'isolement des souches) est recherchée par la méthode statistique de régression implémentée dans le logiciel Path-O-Gen. Le coefficient de détermination ( $R^2$ ) estimé permet d'évaluer la relation entre ces deux variables. L'intersection de la droite de régression linéaire avec l'axe des abscisses donne une approximation de l'âge de l'ancêtre commun le plus récent (TMRCA) de l'échantillon des séquences analysées.



**Figure 41. Phylogénies des quatre sous génogroupes d'EV-A71**

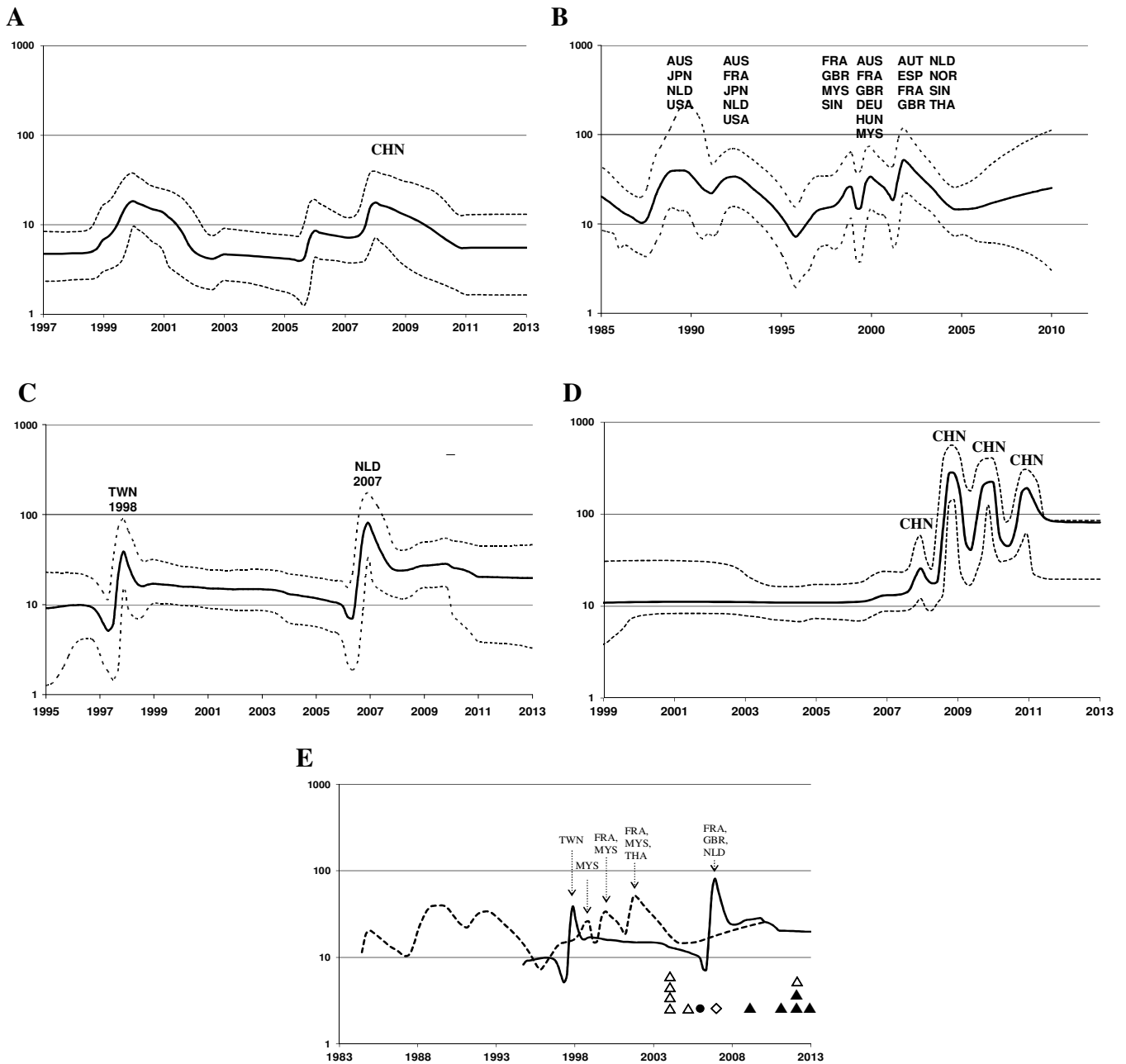
Les chronogrammes ont été inférés avec les jeux de données des séquences 1D<sup>VP1</sup> des sous génogroupes B4/B5 (A), C1 (B), C2 (C) et C4 (D). Les topologies indiquent les relations phylogénétiques des taxons en fonction du temps représenté en abscisses. Les topologies phylogénétiques ont été inférées par une approche Bayésienne utilisant un modèle d'horloge moléculaire assouplie implémentés dans le logiciel BEAST v1.7.5. Par clarté, les noms des taxons ne sont pas indiqués. Les astérisques indiquent les nœuds fiables (probabilité postérieure > 0.90). Les branches sont colorées en fonction de l'origine géographique des taxons : bleu, Europe ; turquoise, Russie ; vert, Amérique et rouge, Asie.

moléculaire et de démographie. Le modèle le plus adéquat pour les quatre jeux des séquences d'EV-A71 est une combinaison du modèle d'évolution moléculaire GTR, du modèle d'horloge moléculaire assouplie dans lequel la variation des taux d'évolution entre les branches suit une loi log normale et du modèle de démographie BSP.

Les séquences incluses dans les jeux de données peuvent être associées à différentes origines géographiques selon le pays dans lequel les souches ont été isolées. La distribution de quatre régions géographiques (Asie, Europe, Amérique et Russie) a été analysée dans les jeux de données et les arbres phylogénétiques inférés par une approche bayésienne [Figure 41]. La distribution des régions géographiques n'est pas uniforme pour les différents jeux de données. Les souches des sous-génogroupes B4, B5 et C4 ont été principalement échantillonnées dans la région Asie-Pacifique et les séquences « européennes » sont rares. L'échantillonnage géographique des souches C1 et C2 est diversifié et comprend un grand nombre de taxons en Europe.

Les arbres MCC estimés pour les populations B4B5, C1, C2 et C4, représentés sous la forme de chronogrammes, révèlent une topologie dite « en escalier » caractérisée par un long tronc auquel se rattachent majoritairement des branches courtes [Figure 42]. Le tronc représente la lignée virale qui persiste au cours du temps alors que les branches indiquent les lignées virales qui subissent des événements d'extinction rapide. Ce profil phylogénétique s'observe typiquement pour les sous-génogroupes C1, C2 et C4. Parmi les nombreuses lignées à courte durée de vie dans les arbres de C1 et C2, des longues branches indiquent l'existence de lignées dont la durée de vie est nettement plus longue mais pour lesquelles il existe un défaut d'échantillonnage conséquent. Ce biais peut être d'ordre temporel dans les régions géographiques déjà échantillonnées ou, ce qui est plus vraisemblable, d'ordre géographique. Le chronogramme des sous-génogroupes B4B5 indique que les souches B4 se sont subdivisées en deux lignées indépendantes : une qui semble s'être éteinte en 2003 et une seconde à l'origine du sous-génogroupe B5.

Les chronogrammes ont permis de dater l'origine des lignées principales [Figure 41]. Les sous-génogroupes B4, C2 et C4 ont émergé à la même période, au milieu des années 1990 ; leurs tMRCA respectifs sont les suivants : 1994,7 (1991,6 – 1996,5), 1993,8 (1992,3 – 1994,7) et 1992,8 (1986,7 – 1996,5). La population C1 est apparue 10 ans plus tôt (tMRCA = 1984,5 (1983,2 – 1985,4)) et la population B5 en 2002,5 (2001,7 – 2003,2).



**Figure 42. Profils de l'évolution démographique (diversité génétique) au cours du temps des sous génogroupes d'EV-A71 B4-B5 (A), C1 (B), C2 (C) et C4 (D)**

La diversité génétique globale au cours du temps a été estimée à partir des séquences du gène  $1D^{VP1}$  pour chaque sous génogroupe d'EV-A71. L'axe des abscisses indique le temps (en années). L'axe des ordonnées représente la diversité génétique ( $\log_{10}N_e(t)$  où  $N_e$  est la taille efficace de la population et  $t$ , le temps de génération). Cette diversité génétique a été estimée par la méthode du Bayesian skyline plot (Drummond, 2005) en utilisant le modèle de substitution nucléotidique GTR et un modèle d'horloge moléculaire relaxée. La valeur moyenne du nombre effectif d'infections est représentée par la ligne continue. L'intervalle de crédibilité à 95% est représenté par les pointillées. Les profils estimés pour les sous génogroupes C1 et C2 ont été superposés dans la figure E et sont représentés respectivement par les lignes en pointillée et pleine. Les pics épidémiques sont numérotés et les principales origines géographiques sont indiquées. Les introductions de lignées du sous génogroupe C4 en Europe et en Russie sont indiquées respectivement avec les triangles ouverts et pleins. Les récents évènements d'introduction du sous génogroupe B5 au Danemark (Badran et al, 2010) et en France (Mirand et al, 2015) sont montrés avec des losanges. La chaîne de transmission de souches du C4 de l'Europe vers le Canada est indiquée par des triangles connectés.

### C. Histoire démographique des populations d'EV-A71 en Europe et en Asie

L'histoire démographique des populations virales a été inférée conjointement aux phylogénies présentées ci-dessus. Les profils démographiques montrent des variations entre les cinq sous-génogroupes d'EV-A71 [Figure 42].

Le profil du sous-génogroupe B4 indique une augmentation importante de la taille de la population entre 1999 et 2001, période qui correspond à la survenue d'épidémies rapportées au Japon, à Singapour et Taiwan en 2000 [Figure 42A]. L'apparition du sous-génogroupe B5 en 2003 ne coïncide pas avec une augmentation de la taille de la population. De 2006 à 2009, la population B5 augmente ; des épidémies ont été rapportées en Malaisie, en Thaïlande, à Singapour et à Taiwan pendant cette période.

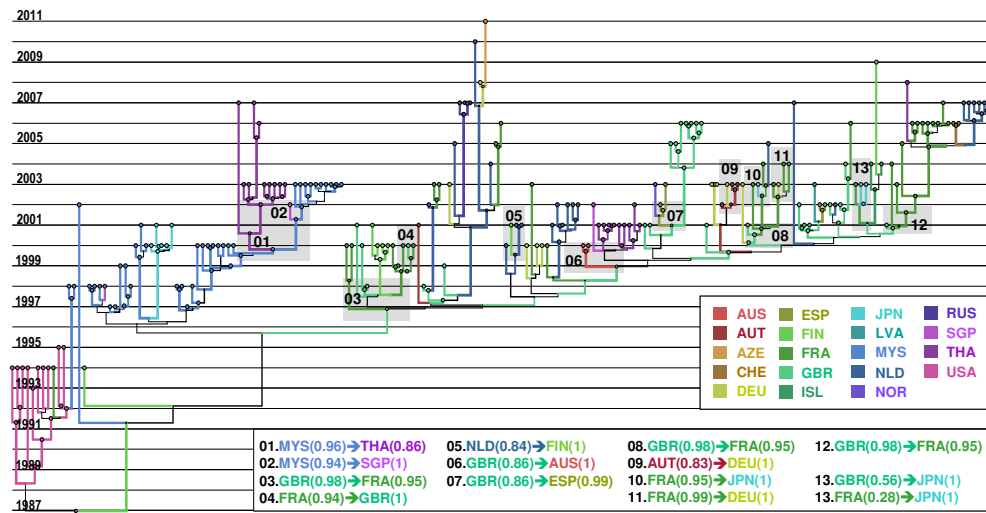
Le profil démographique du sous-génogroupe C1 montre une succession de faibles variations de la taille de la population virale au cours du temps, ce qui est en accord avec la faible activité épidémique du virus [Figure 42B].

L'évolution démographique du sous-génogroupe C2 est caractérisée par deux augmentations brusques de la taille de sa population [Figure 42C]. Le premier pic se situe en 1998, ce qui correspond à la survenue d'une importante épidémie à Taiwan. Par la suite, la taille de la population diminue jusqu'à la survenue d'un second pic en 2007 d'une amplitude plus importante que le premier. Il est la conséquence d'une épidémie survenue aux Pays-Bas en 2007. Après 2007, la démographie du sous-génogroupe C2 demeure stable.

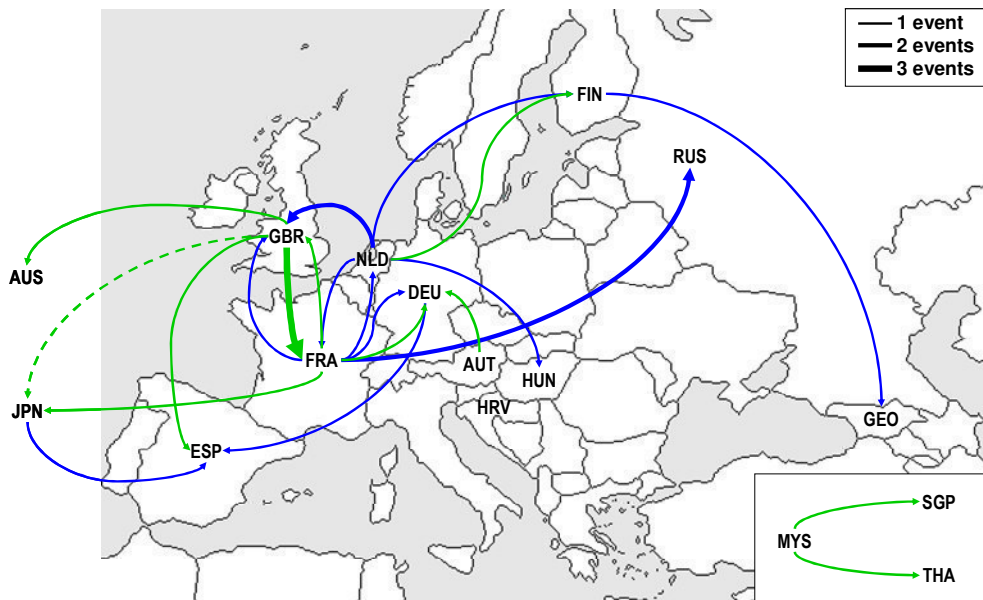
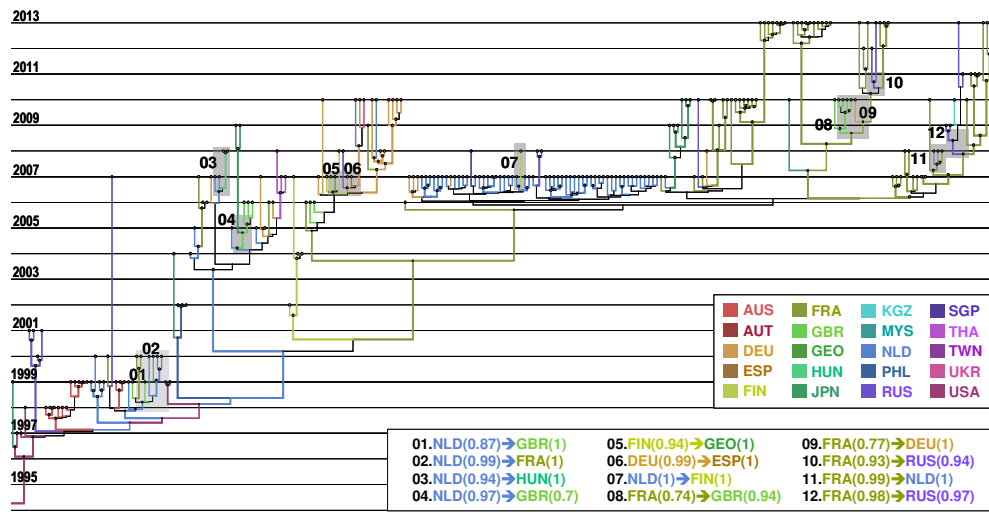
Le profil du sous-génogroupe C4 montre la persistance d'une population de petite taille pendant les 10 premières années de circulation de ce virus, probablement due à un défaut d'échantillonnage durant cette période [Figure 42D]. Quatre augmentations démographiques ont été estimées : elles correspondent à des épidémies annuelles de très grande ampleur rapportées en Chine. La première en 2007 est caractérisée une faible incidence, comparée aux trois suivantes, en 2008, 2009 et 2010.

La superposition des profils démographiques des sous-génogroupes C1 et C2 [Figure 42E] montre que les deux pics d'augmentation de la taille de la population C2 sont survenus alors que celle de la population C1 était faible. Le pic estimé pour l'épidémie de Taiwan en 1998 a eu lieu alors que la population C1 présentait globalement une faible croissance démographique. Après 1999, la population C1 présente plusieurs augmentations démographiques (de moindre ampleur que les deux pics de C2) pendant une longue période de stagnation démographique de la population C2. De 2004 à 2006, la taille des populations virales C1 et C2 stagne alors que des infections sporadiques causées par des souches C4 sont rapportées en Europe puis au Canada. En 2007, la population C1 est faible lorsque survient

A

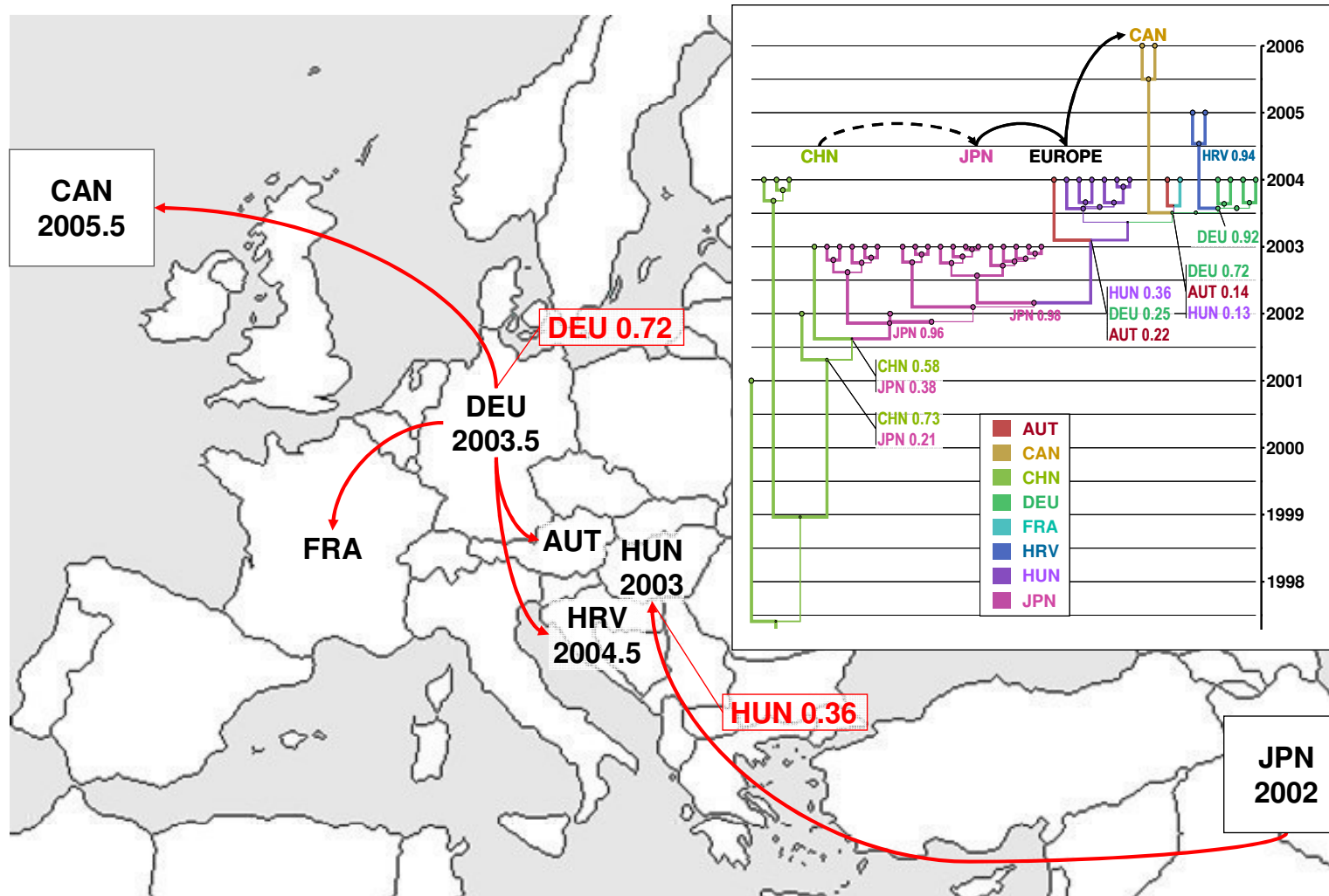


B



### Figure 43. Analyse bayésienne de la distribution spatiale et temporelle des souches des sous-génogroupes C1 et C2 d'EV-A71

Les chronogrammes C1 (A) et C2 (B) ont été inférés avec une approche bayésienne, un modèle d'horloge moléculaire relaxée et un modèle de caractères discrets phylogéographiques implémentés dans le logiciel BEAST v1.7.5. Les noms des taxons ne sont pas indiqués pour rendre les arbres lisibles. L'épaisseur des branches est proportionnelle à la fiabilité de l'embranchement (probabilité postérieure). Les branches et les nœuds sont colorés en fonction de l'origine géographique la plus probable estimée par l'analyse phylogéographique. Les nombres indiquent les événements de transports les plus probables. Ils sont représentés sur la carte selon le code couleur, C1 en vert et C2 en bleu (C). Les pays sont désignés par la nomenclature à trois lettres du code iso 3166.



**Figure 44. Analyse bayésienne de la distribution spatiale et temporelle d'un cluster de souches du sous génogroupe C4 d'EV-A71**

Le chronogramme a été inféré avec une approche bayésienne, un modèle d'horloge moléculaire relaxée et un modèle de caractères discrets phylogéographiques implémentés dans le logiciel BEAST v1.7.5. Par clarté, les noms des taxons ne sont pas indiqués. L'épaisseur des branches est proportionnelle à la fiabilité de l'embranchement (probabilité postérieure). Les branches et les nœuds sont colorés en fonction de l'origine géographique la plus probable estimée par l'analyse phylogéographique. Les événements de transports de virus les plus probables sont représentés sur la carte à gauche. Les pays sont désignés par la nomenclature à trois lettres du code iso 3166.



aux Pays-Bas l'épidémie causée par des souches C2 (pic n°2). Ces caractéristiques démographiques évoquent une interrelation entre les deux populations virales à une échelle géographique globale.

#### **D. Phylogéographie des populations européennes d'EV-A71**

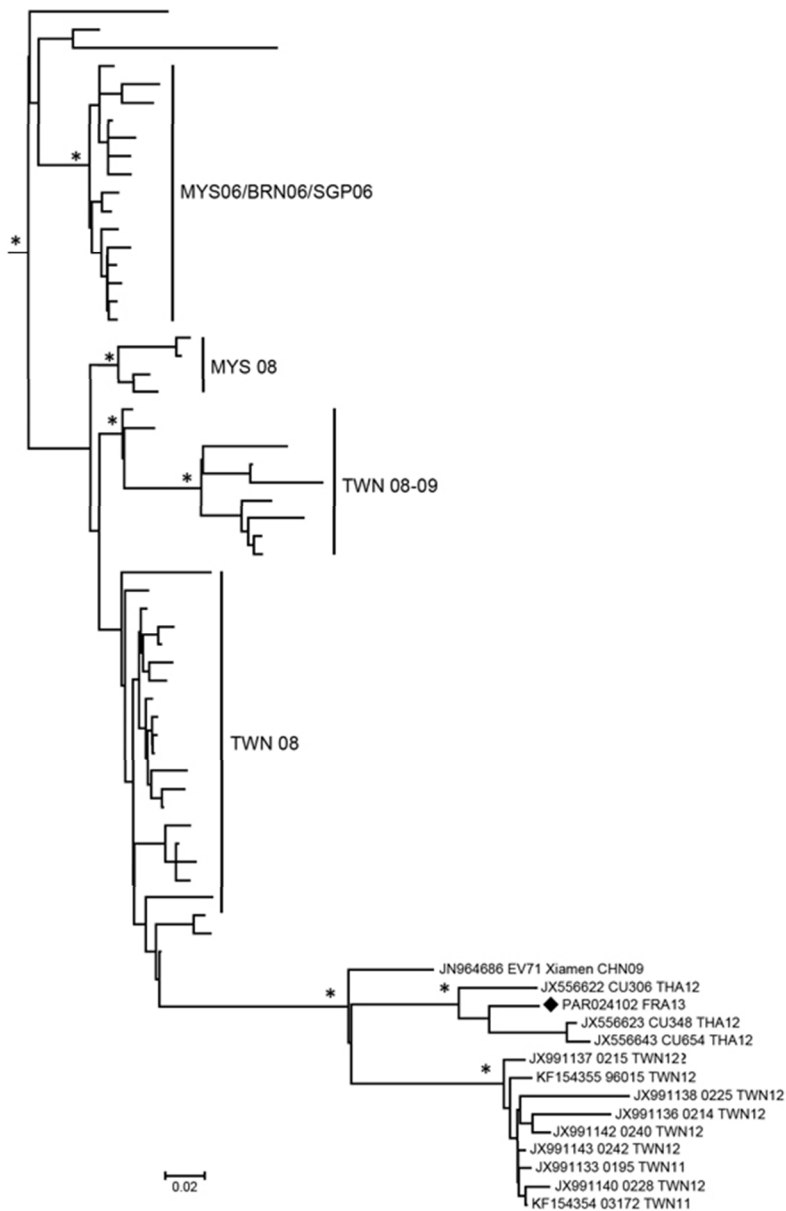
La probabilité de transmission des souches d'EV-A71 entre les pays échantillonnés a été estimée en utilisant un modèle de diffusion de caractères discrets, appliqué aux phylogénies décrites ci-dessus. Un évènement de transport de souches virales entre deux pays a été considéré comme probable entre deux nœuds reliés directement par une branche, lorsque les trois caractéristiques phylogénétiques suivantes étaient réunies [Figure 43] :

- (1) les nœuds présentent une densité de probabilité postérieure  $> 0,9$
- (2) la densité de probabilité de l'origine géographique des nœuds  $> 0,7$
- (3) la différence entre les tMRCA estimés pour chaque nœud est d'un an ou moins.

Avec ces critères, 13 évènements de transmission ont été identifiés dans la topologie de l'arbre inféré pour le sous-génogroupe C1 [Figures 43A, 43C]. Huit d'entre eux (n° 03, 04, 05, 07, 08, 09, 11 et 12) représentent une transmission entre pays européens. La Grande-Bretagne a été la source de trois transports de souches vers la France (n° 03, 08 et 12 ; pp=0,95) et d'un évènement vers l'Espagne (n° 07 ; pp=0,99). La France a été la source de deux évènements de transmission : l'un vers l'Allemagne (n° 11 ; pp=0,99), l'autre vers la Grande Bretagne (n° 04 ; pp=0,94). D'autres transports de souches ont eu lieu : de l'Autriche vers l'Allemagne et des Pays-Bas vers la Finlande. Trois évènements (n° 06, 10 et 13) ont été estimés entre des zones géographiques distantes : Europe (Allemagne et France) et Asie (Malaisie et Japon). Toutefois, la source européenne de l'évènement n° 13 demeure incertaine (probabilité d'une origine en France (p=0,28), en Grande Bretagne (p=0,56). Enfin, deux transports de souches ont été estimés en Asie, avec comme source la plus probable la Malaisie et comme pays d'arrivée la Thaïlande (évènement n° 01) et Singapour (évènement n° 02).

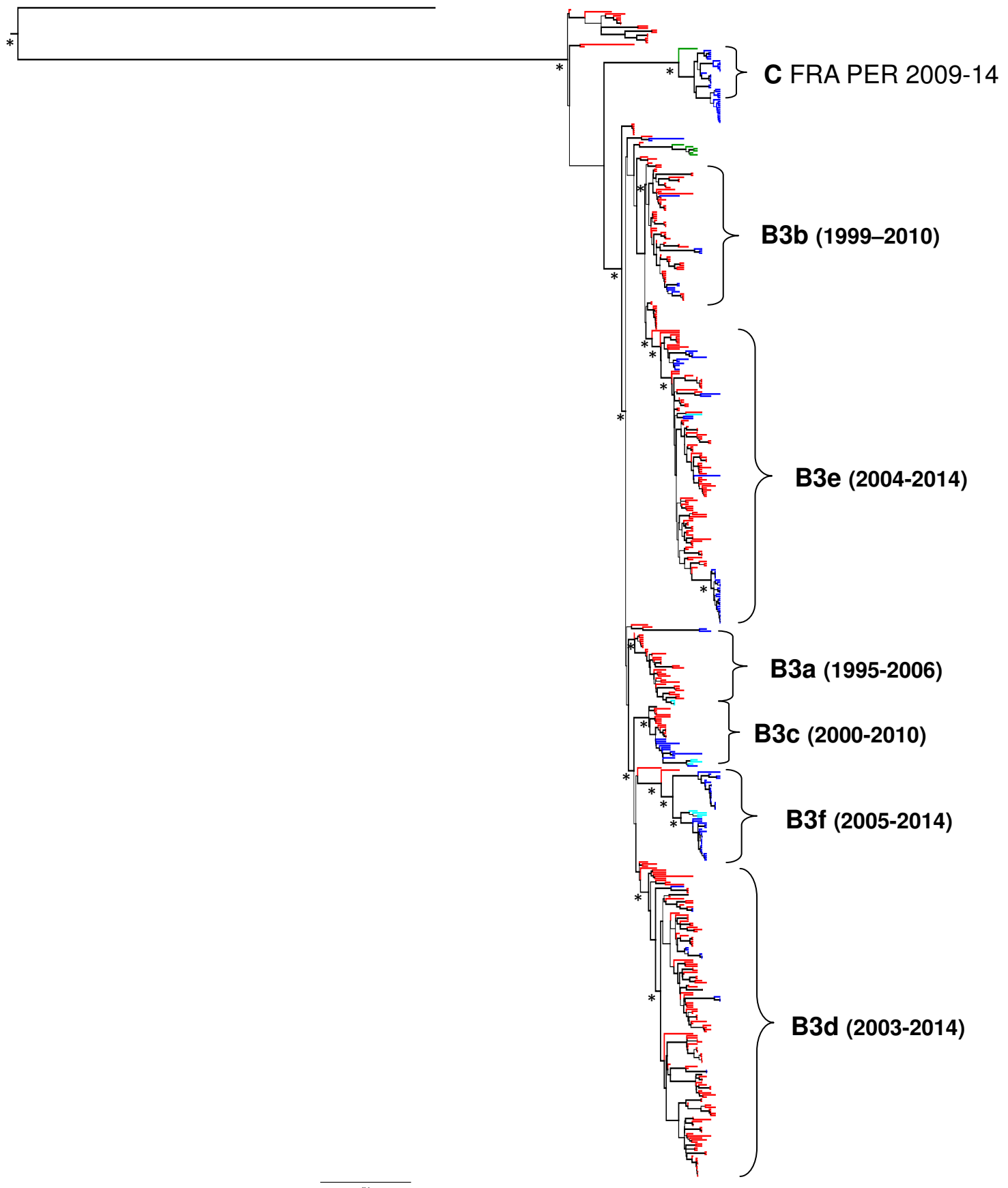
La généalogie du sous-génogroupe C2 a permis d'estimer 12 transmissions entre des pays en Europe occidentale et orientale [Figures 43B, 43C]. Une majorité d'évènements de transport ont comme pays sources les Pays-Bas (n=5 ; pp=0,87 – 1,0) et la France (n=5 ; pp=0,74 – 0,98). Trois transmissions virales sur de longues distances ont été estimées : une entre la Finlande et la Géorgie et deux entre la France et la Russie.

Concernant le sous-génogroupe C4, la recherche de transmissions de souches a été focalisée principalement sur le groupe monophylétique représenté en gris sur le



**Figure 45. Phylogramme de l'EV-A71 B4/B5 inféré à partir de 274 séquences du gene 1D<sup>VP1</sup> partiel.**

Le losange noir indique la souche isolée en France en 2013. Les topologies phylogénétiques ont été inférées par une approche Bayésienne utilisant un modèle d'horloge moléculaire assouplie implémentés dans le logiciel BEAST v1.7.5. Par clarté, les noms des taxons ne sont pas indiqués. Les astérisques indiquent les nœuds fiables (probabilité postérieure > 0.90). Les pays sont désignés par la nomenclature à trois lettres du code iso 3166.



**Figure 46. Généalogie du CVA16**

Le chronogramme a été inféré avec le jeu de données du locus 1D<sup>VP1</sup> du CV-A16. La topologie indique les relations phylogénétiques des taxons en fonction du temps. Cette représentation phylogénétique a été inférée avec une approche bayésienne et un modèle d’horloge moléculaire relaxée implémentés dans le logiciel BEAST v1.7.5. Les noms des taxons ne sont pas indiqués pour rendre l’arbre plus lisible. Les étoiles indiquent les nœuds fiables (probabilité postérieure > 0.90). Les branches sont colorées en fonction de l’origine géographique des taxons : bleu, Europe ; turquoise, Russie ; vert, Amérique et rouge, Asie.

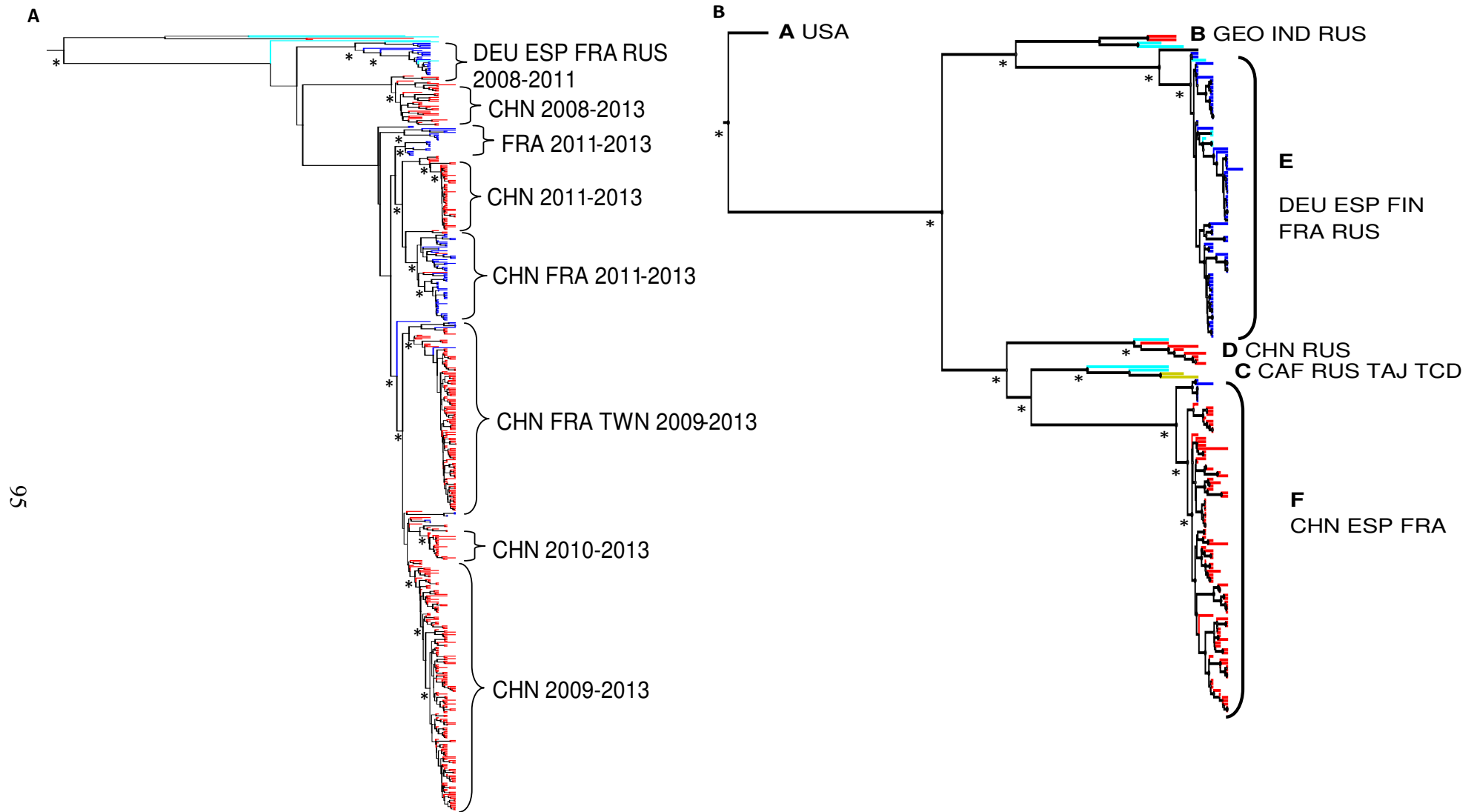
chronogramme de la Figure 41D. Ce groupe est composé de 41 taxons issus des pays suivants : Allemagne, Autriche, Canada, Chine, Croatie, France, Hongrie et Japon. Une analyse phylogénétique a été effectuée pour ce groupe et les probabilités de l'origine géographique sont indiquées pour les nœuds fiables [Figure 44]. Le profil phylogéographique suggère que la souche virale a circulé entre la Chine et le Japon entre 2001 et 2002 et qu'elle a été transportée en Europe au début de l'année 2003, toutefois le pays précis qui a été touché en premier n'a pas pu être identifié (Hongrie,  $pp=0,36$ ). La même souche a ensuite été rapportée dans plusieurs pays européens (Allemagne, Autriche, Croatie, France, Hongrie) en 2004. Un nœud de la phylogénie suggère l'Allemagne ( $pp=0,72$ ) comme source probable de la souche rapportée en Croatie en 2005 et au Canada où le virus a été isolé en 2006.

Une analyse a été réalisée spécifiquement pour étudier l'histoire évolutive de la première souche du sous génogroupe B5 qui a été isolée en France en 2013. Le profil phylogénétique et la phylogéographie indique que ce virus a été introduit en France depuis l'Asie (probablement la Thaïlande) en 2010,6 environ par datation moléculaire [Figure 45].

### **3) Epidémiologie moléculaire des coxsackievirus A**

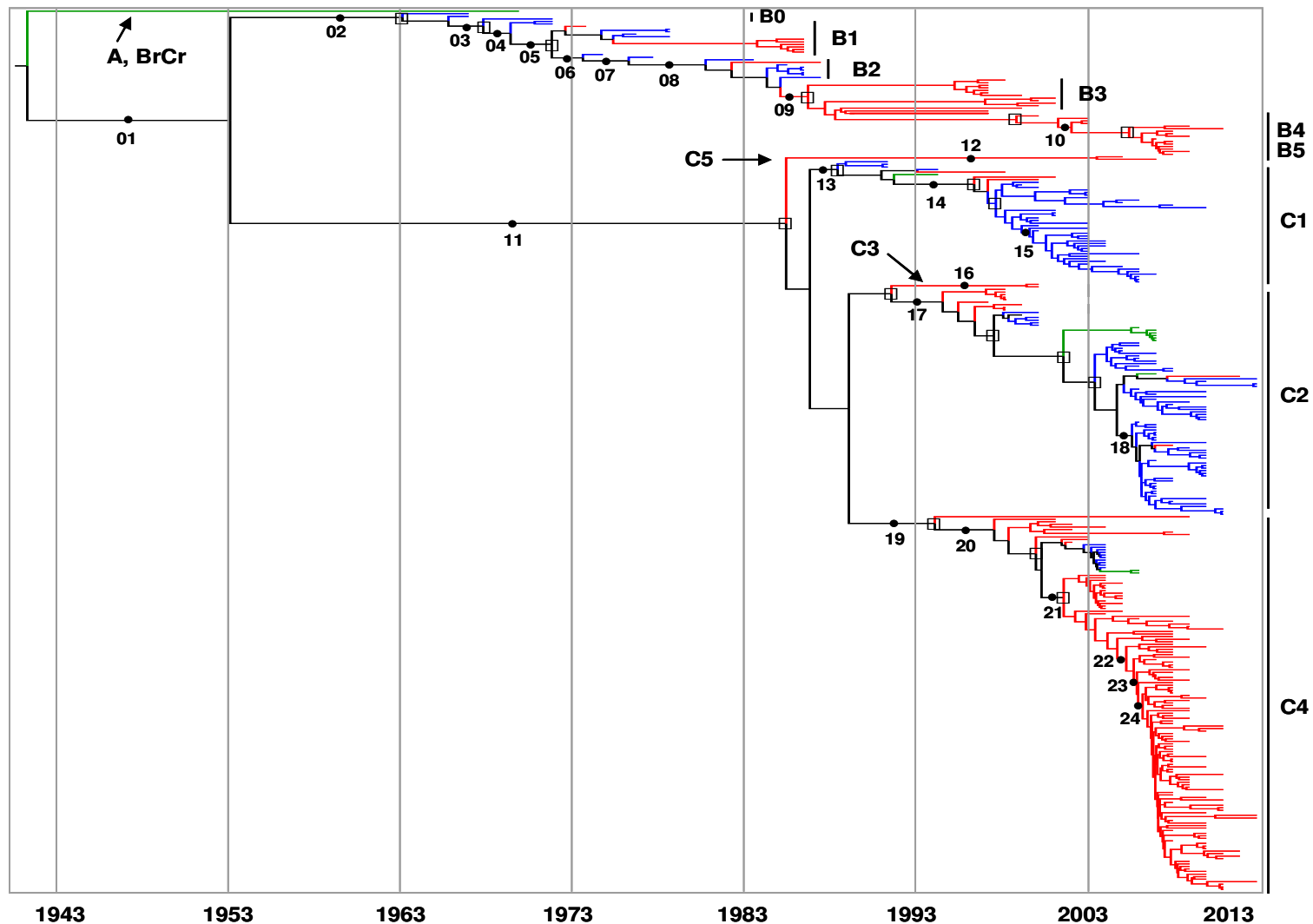
La généalogie du CV-A16 se compose d'une multitude de clades dont la fiabilité est validée par une probabilité postérieure supérieure à 0,90 [Figure 46]. Le sérotype CV-A16 comprend 3 génogroupes A, B et C et différents sous-génogroupes B1, B2 et B3. La souche prototype G-10, isolée 1949, compose à elle seule le génogroupe A. Les populations B1 et B2 comprennent des souches anciennes isolées en Asie (Japon) entre 1981 et 2000. Le sous-génogroupe B3 se décompose en six lignées distinctes de souches qui co-circulent entre l'Europe et l'Asie pendant la période 1995–2014. Le génogroupe C est un clade récent, issu de souches isolées en France entre 2011 et 2014 et d'une souche isolée au Pérou en 2009. L'analyse phylogéographique révèle quelques événements de transmissions, en majorité « locaux » (en Asie ou en Europe) et quelques événements de transmission transcontinentaux entre la France et la Chine.

La généalogie du CV-A6 se décompose en 8 clades monophylétiques ( $pp>0,90$ ) [Figure 47]. Ces clades comprennent des souches isolées principalement en Chine ou en France. De nombreux nœuds ne sont pas fiables au niveau du tronc principal, probablement en raison d'un échantillonnage trop limité par les deux principales origines géographiques. Les clades indiquent la circulation de plusieurs populations virales ; ce qui est en accord avec les données épidémiologiques rapportées pour ce sérotype concernant les épidémies de



**Figure 47. Généalogies du CVA6 (A) et du CV-A10 (B)**

Les chronogrammes ont été inférés avec les jeux de données du locus  $1D^{VP1}$ . Les topologies indiquent les relations phylogénétiques des taxons en fonction du temps représenté en abscisses. Ces représentations phylogénétiques ont été inférées avec une approche Bayésienne et un modèle d'horloge moléculaire assouplie implémentés dans le logiciel BEAST v1.7.5. Les noms des taxons ne sont pas indiqués. Les étoiles indiquent les nœuds fiables (probabilité postérieure > 0.90). Les branches sont colorées en fonction de l'origine géographique des taxons : bleu, Europe ; turquoise, Russie ; vert, Amérique ; rouge, Asie et jaune, Afrique.



**Figure 48. Généalogie de l'EV-A71 inférée avec les séquences génomiques P1**

Le profil évolutif de 344 souches d'EV-A71 échantillonnées entre 1962 et 2013 a été reconstruit avec la région génomique P1 (2586 nucléotides). La topologie phylogénétique a été reconstruite avec une méthode Bayésienne utilisant un modèle d'horloge moléculaire relaxée implémentés dans le logiciel BEAST v1.7.5. Les nœuds fiables (probabilité postérieure > 0.90) sur le tronc principal sont représentés par les carrés vides. Les branches sont colorées en fonction de l'origine géographique des taxons : bleu, Europe ; vert, Amérique et rouge, Asie. Par clarté, les noms des taxons ne sont pas indiqués. Les nombres indiquent les branches où les substitutions non synonymes ont été inférées avec la méthode Ancestral State Reconstruction (ASR).

maladie pied-main-bouche. L'analyse phylogéographique du CV-A6 n'a pas révélé d'informations importantes, en raison du faible échantillonnage.

La généalogie du CV-A10 se caractérise par deux clades majoritaires (E et F) dont les souches circulent en Europe et en Asie entre 2008 et 2014 [Figure 47]. Trois autres clades comprennent des souches d'origines géographiques hétérogènes (Russie, Inde, Chine, Afrique). Le clade E est composé de souches isolées en France (n=70/85), réparties en 9 lignées qui ont co-circulé en 2010 (5 lignées) et 2012 (4 lignées). Le clade F contient une majorité de souches isolées en Chine (n=91 souches ; 2008–2012) et une lignée de 7 souches isolées en Europe entre 2008 et 2010 (Espagne et France). L'analyse phylogéographique permet de décrire 5 évènements de transmission entre différents pays d'Europe occidentale (Finlande, France, Espagne) et d'Europe orientale (Russie) [données non montrées].

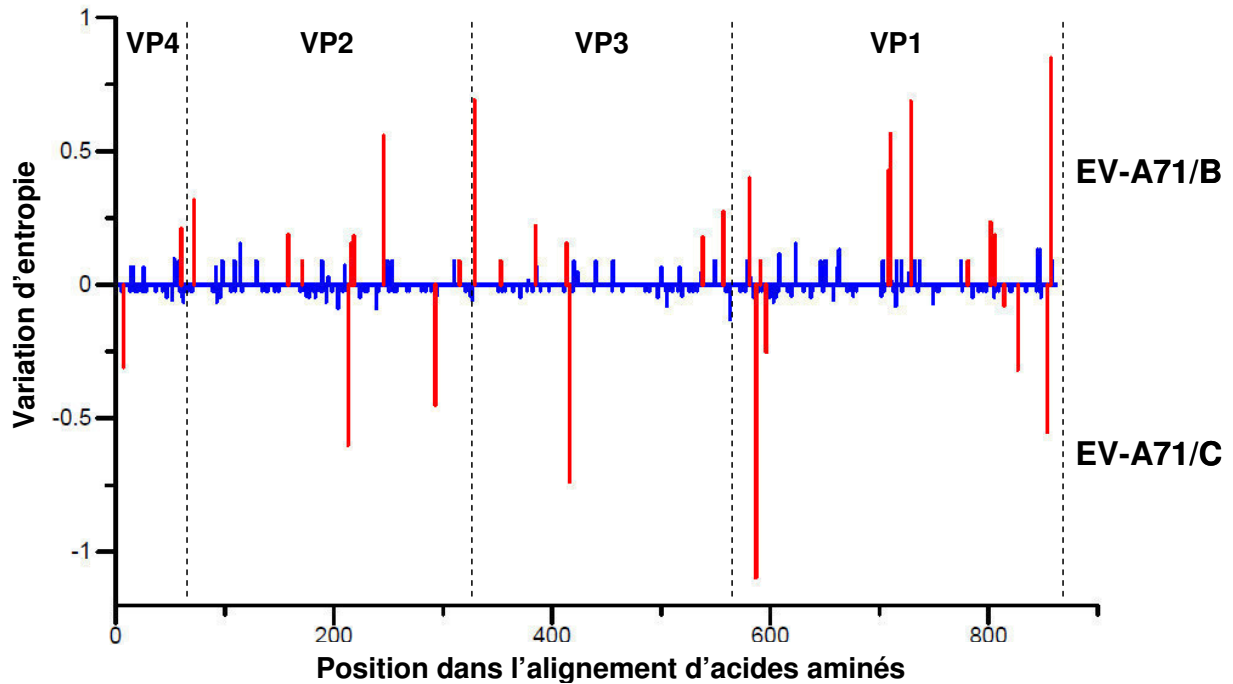
## II. Evolution moléculaire de la capsid de l'EV-A71

### 1) Phylogénie moléculaire de l'EV-A71 établie avec les séquences de la région P1

La phylogénie des sous-génogroupes d'EV-A71 a été estimée avec un échantillon de 344 séquences de la région P1 [Figure 48]. La racine du chronogramme a été datée en 1937,3 (1913,7–1956,3) période à laquelle s'est produite la séparation entre le génogroupe A et la lignée des génogroupes B et C. Le dernier ancêtre commun des génogroupes B et C a été daté en 1953 (1939,5–1960,8).

La généalogie des souches du génogroupe B se caractérise par un long tronc duquel partent les lignées correspondant aux sous-génogroupes B0 à B5. Les lignées se remplacent séquentiellement et présentent une faible diversité génétique. La topologie typiquement « en escalier » implique un turnover rapide des souches du génogroupe B et l'extinction des lignées correspondant aux anciens sous-génogroupes.

*A contrario*, le génogroupe C se caractérise par une diversification (ramification) des lignées correspondant aux sous-génogroupes C1 à C5 au cours d'une courte période (1984,9–1988,7). Le profil phylogénétique en étoile indique que les populations virales étaient en co-circulation pendant la période échantillonnée, mais pas uniquement dans les deux zones géographiques (Asie et Europe). Le sous-génogroupe C5 s'est individualisé de l'ancêtre commun du génogroupe C en 1984,9 (1978,8–1987,6). L'embranchement de l'origine du sous-génogroupe C1 n'est pas fiable (pp << 0,9) ; ceux des ancêtres communs propres aux sous-génogroupes C2 et C4 présentent une densité de probabilité postérieure proche de 0.9



**Figure 49. Variabilité en acides aminés des protéines de capsid de l'EV-A71**

La variabilité a été estimée par une analyse d'entropie de chaque position d'acide aminé au sein des séquences des génogroupes B et C. Les différences d'entropies sont calculées entre les séquences protéiques des deux génogroupes. Les différences significatives d'entropie ( $P \leq 0.05$ ) sont représentées en rouges. Les limites des protéines de capsid sont représentées par les pointillés (protéine VP4, positions 1 à 68 ; VP2, 69 à 322 ; VP3, 323 à 564 et VP1, 565 à 861).



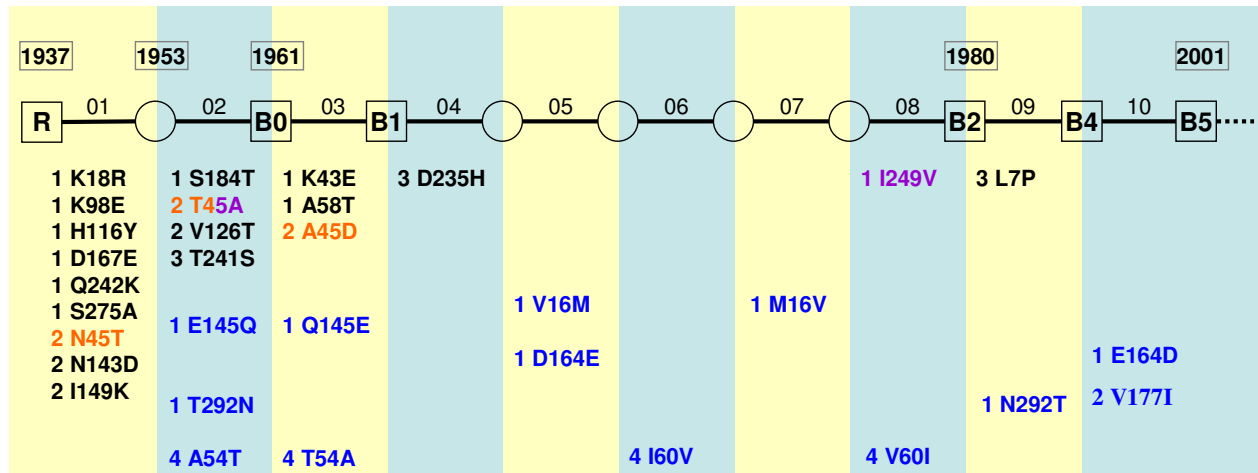
**Tableau 12** : Acides aminés observés au niveau de la région P1 des souches d'EV-A71.

Position <sup>a</sup>	Protéine	Position au sein de la protéine	Motif de la protéine VP1 <sup>b</sup>	Acide aminé pour le génogroupe B d'EV-A71	Acide aminé pour le génogroupe C d'EV-A71
7	VP4	7		T (n=55) ; A (n=1)	T (n=248) ; A (n=39)
60	VP4	60		I (n=53) ; V (n=3)	I (n=287)
72	VP2	3		S (n=50) ; T (n=6)	S (n=286) ; T (n=1)
158	VP2	89	Boucle C-D	T (n=53) ; I (n=3)	T (n=286) ; I (n=1)
171	VP2	102	Boucle C-D	Y (n=55) ; N (n=1)	Y (n=287)
213	VP2	144	Boucle E-F	S (n=56)	S (n=204) ; T (n=83)
216	VP2	147	Boucle E-F	P (n=54) ; H (n=2)	P (n=287)
218	VP2	149	Boucle E-F Hélice $\alpha 2^*$	K (n=52) ; I (n=4)	K (n=283) ; M (n=4)
246	VP2	177	Boucle E-F Hélice $\eta 6$	V (n=44) ; I (n=11) ; M (n=1)	V (n=286) ; I (n=1)
293	VP2	224	Boucle H-I	F (n=56)	F (n=240) ; Y (n=47)
315	VP2	246	Boucle H-I	A (n=55) ; G (n=1)	A (n=287)
329	VP3	7	N-terminal	P (n=30) ; L (n=26)	L (n=287)
353	VP3	31	N-terminal	P (n=55) ; L (n=1)	P (n=287)
385	VP3	63	Hélice H1	A (n=53) ; V (n=2) ; D (n=1)	A (n=286) ; V (n=1)
413	VP3	91	Boucle C-D	P (n=54) ; L (n=2)	P (n=287)
414	VP3	92	Boucle C-D	G (n=54) ; E (n=2)	G (n=287)
416	VP3	94	Boucle C-D	D (n=56)	N (n=178) ; S (n=104) ; D (n=5)
538	VP3	216	Brin $\beta 1$	I (n=54) ; L (n=1) ; M (n=1)	I (n=287)
557	VP3	235	C-terminal	H (n=52) ; D (n=3) ; Y (n=1)	D (n=286) ; E (n=1)
581	VP1	16	N-terminal	V (n=47) ; M (n=9)	V (n=285) ; M (n=2)
587	VP1	22	N-terminal	Q (n=56)	Q (n=101) ; H (n=95) ; R (n=91)
591	VP1	26	N-terminal	A (n=55) ; T (n=1)	A (n=287)
596	VP1	31	N-terminal	N (n=56)	N (n=267) ; D (n=20)
708	VP1	143	Boucle D-E	T (n=49) ; S (n=6) ; A (n=1)	T (n=287)
710	VP1	145	Boucle D-E	E (n=38) ; G (n=8) ; Q (n=10)	E (n=266) ; Q (n=18) ; G (n=3)
729	VP1	164	Boucle E-F	E (n=31) ; D (n=25)	D (n=287)
781	VP1	215	Boucle G-H*	Q (n=55) ; H (n=1)	Q (n=287)
802	VP1	237	Brin $\beta H$	T (n=52) ; N (n=4)	T (n=286) ; N (n=1)
806	VP1	241	Boucle H-I	S (n=53) ; L (n=3)	S (n=286) ; L (n=1)
814	VP1	249	Brin $\beta 1$	V (n=39) ; I (n=17)	V (n=148) ; I (n=139)
827	VP1	262	Brin $\beta 1$	I (n=56)	I (n=257) ; V (n=30)
854	VP1	289	C-terminal	T (n=55) ; A (n=1)	A (n=202) ; T (n=83) ; V (n=2)
857	VP1	292	C-terminal	T (n=31) ; N (n=23) ; S (n=2)	T (n=285) ; A (n=2)

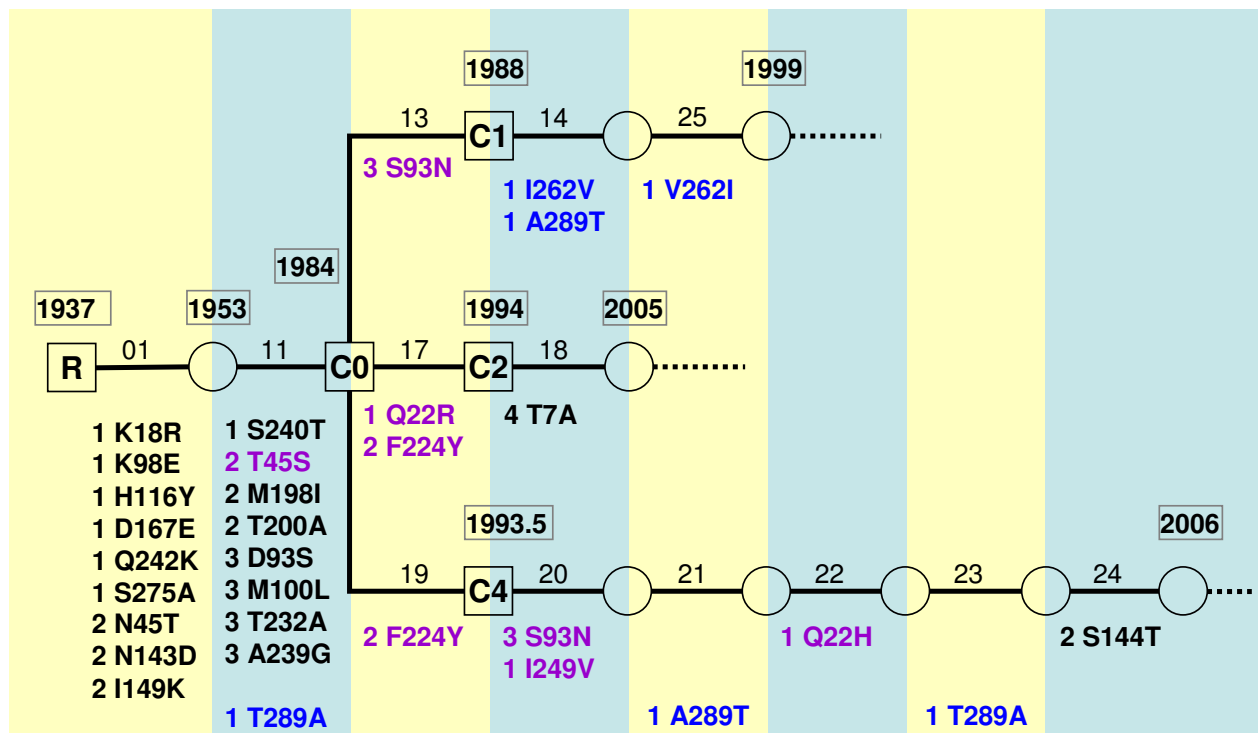
<sup>a</sup> Position du codon à l'intérieur de la région P1.

<sup>b</sup> Localisation des acides aminés au niveau des motifs de la structure secondaire de la protéine de capsid VP1 de l'EV-71 décrite dans un travail récent (Wang W *et al.*, 2012) Les étoiles indiquent les positions situées au sein d'un épitope déjà mis en évidence.

A



B



**Figure 50. Représentation schématique des trajets évolutifs reconstruits pour les génogroupes B et C de l'EV-A71**

La généalogie représentée dans la Figure 48 a été simplifiée afin de représenter uniquement le tronc principal des lignées importantes (génogroupes, sous-génogroupes) des génogroupes B (A) et C (B). Les branches sont numérotées selon la généalogie originale et les substitutions en acides aminés sont également représentées. Les nœuds ancestraux des lignées connues sont représentés par les carrés. Les nœuds ancestraux intermédiaires sont indiqués par des cercles. Les changements en acides aminés sont indiqués de la manière suivante : numéro de la protéine de capsid, résidu initial, position au niveau de la protéine de capsid et résidu final. Le code couleur des substitutions non synonymes indique les changements séquentiels (orange), les changements réverses (bleu) et les changements parallèles (violet).

( $pp = 0,88$ ). Ces caractéristiques ne semblent pas être un artefact de la reconstruction phylogénétique mais suggère que des sous-populations de C1 sont à l'origine des autres sous-génogroupes.

## **2) Variation d'entropie dans les séquences protéiques P1 de l'EV-A71**

Pour comprendre l'évolution moléculaire de l'EV-A71 au cours du temps, c'est-à-dire au cours de la transmission des souches, la diversité a été analysée dans les quatre protéines de capsid et à deux niveaux : la diversité protéique au niveau des feuilles de l'arbre et celle qui se manifeste entre la racine et les branches terminales. La première a été étudiée en déterminant l'entropie, la seconde avec la méthode ASR (cf. paragraphe 3 ci-dessous).

La variation d'entropie a été calculée à chaque site en acide aminé entre les séquences P1 des génogroupes B et C pour estimer la diversité relative entre les deux populations virales. Les profils d'entropie montrent que chacune des quatre protéines de capsid présente un polymorphisme protéique propre à chacun des deux génogroupes [Figure 49]. Des variations d'acides aminés ont été estimées de façon fiable pour 33 résidus dont 24 sont polymorphes chez le génogroupe B et 9 chez le génogroupe C. Le nombre total de sites polymorphes est plus élevé dans la protéine VP1 ( $n=14$ ) que les autres protéines de capsid (VP2,  $n=9$  ; VP3,  $n=8$  ; VP4,  $n=2$ ). La diversité en acides aminés à chaque site et leur localisation dans les motifs de la structure secondaire sont présentées dans le Tableau 12. Près de la moitié des sites polymorphes en acides aminés ( $16/33$ ) sont localisées dans des boucles, les segments protéiques qui relient les feuillets  $\beta$ .

## **3) Estimation des substitutions non synonymes dans la généalogie de l'EV-A71**

Le processus évolutif qui a généré la diversité protéique dans les taxons a été analysé parmi les lignées de l'arbre reconstruit pour la région P1 de l'EV-A71. Les substitutions non synonymes ont été estimées le long de la généalogie par la méthode ASR. L'analyse ASR a été centrée sur les substitutions inférées le long du tronc principal et des branches qui conduisent aux lignées majeures (sous-génogroupes) ; ces substitutions représentent les changements en acides aminés conservés dans les lignées d'EV-A71 qui persistent au cours du temps (contrairement aux lignées terminales).

Cinquante-six substitutions non synonymes ont été inférées sur 24 branches du tronc principal de la généalogie P1 de l'EV-A71 [Figure 50]. Les substitutions se répartissent entre les protéines de la façon suivante : VP1 ( $n=30$  ; 53,6 %), VP2 ( $n=12$  ; 21,4 %), VP3 ( $n=9$  ;

**Tableau 13.** Comparaison des contraintes sélectives subies par les protéines de capsid entre les sous génogroupes d'EV-A71.

Sous génogroupe	SLAC	FEL		IFEL		FUBAR		MEME	PRIME
	dN/dS <sup>a</sup>	NSS <sup>b</sup>	PSS <sup>c</sup>	NSS <sup>b</sup>	PSS <sup>c</sup>	NSS <sup>b</sup>	PSS <sup>c</sup>	Site soumis à une sélection positive épisodique <sup>d</sup>	Site soumis à des modifications biochimiques
EV-A71/B	0.036 (0.030 – 0.041)	68,2 %	<b>663</b> (3) <b>710</b> (10)	57,7 %	<b>710</b> (4,07)	87,3 %	<b>663</b> (1) <b>710</b> (1)	14, 35, 207, 210, 244, 248, 307, 557, 581, 646, <b>663</b> , <b>710</b> , 755, 802, 806, 845	710 (n=1)
EV-A71/C1	0,023 (0,018 – 0,029)	51,9 %	<b>ND</b>	23,3 %	<b>ND</b>	84,4 %	<b>ND</b>	17, 61, 71, 354, 648, <b>710</b> , 749, 807	71 (n=1)
EV-A71/C2	0,025 (0,020 – 0,030)	64,4 %	<b>ND</b>	38,3 %	<b>ND</b>	89,1 %	/	218, 510, 529, 658, 702, <b>710</b> , 721	721 (n=2)
EV-A71/C4	0,041 (0,036 – 0,046)	89,9 %	416 (0,98) <b>663</b> (4,32) <b>710</b> (2,05) 814 (0,25)	49,3 %	416 (1,16)	93,3 %	416 (0,93) <b>663</b> (1) <b>710</b> (0,96)	91, 95, 96, 169, 198, 241, 258, 500, 515, 584, 599, <b>663</b> , <b>710</b> , 845	500 (n=1) 663 (n=1)

<sup>a</sup> Rapport du taux de substitutions non-synonymes sur le taux de synonymes (dN / dS) estimé par la méthode SLAC.

<sup>b</sup> Pourcentage du nombre de sites sélectionnés négativement (NSS) estimé par les méthodes FEL, IFEL et FUBAR.

<sup>c</sup> Position des sites sélectionnés positivement (PSS) estimés par les méthodes FEL, IFEL et FUBAR. Le rapport dN/dS normalisé figure entre parenthèses pour les analyses FEL et IFEL. Pour la méthode FUBAR, les probabilités postérieures sont indiquées entre les parenthèses.

<sup>d</sup> Position des sites soumis à une sélection positive épisodique estimés par la méthode du MEME et qui sont fiables (p-value < 0.05).

<sup>e</sup> Position des sites sélectionnés positivement dont les substitutions entraînent des modifications biochimiques, estimés par la méthode PRIME. Le nombre de modifications biochimiques au site d'intérêt est indiqué entre parenthèses.

**ND**, Non déterminé.

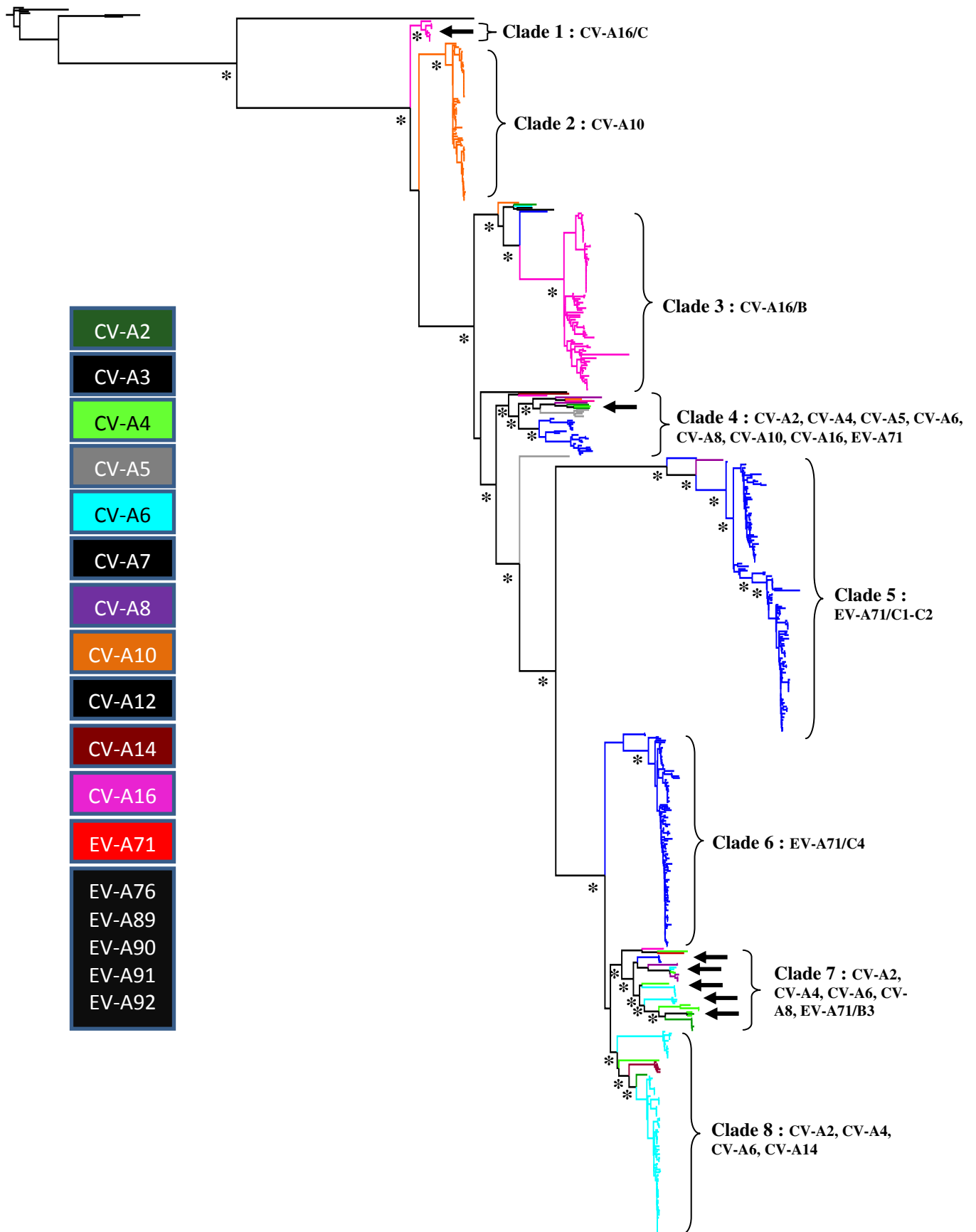
16,1 %) et VP4 (n=3 ; 8,9 %). Un groupe de 13 positions de codons (54, 60, 114, 293, 416, 581, 587, 710, 729, 814, 827, 854 et 857) a subi au moins deux substitutions non synonymes dans la généalogie ; ces substitutions peuvent être des changements réverses, parallèles ou séquentiels d'acides aminés.

L'histoire évolutive du génogroupe B est caractérisée par un plus grand nombre de changements réverses (n=6) que le génogroupe C (n=3) [Figure 50]. Des changements parallèles d'acides aminés ont été observés pendant l'histoire évolutive de l'EV-A71 entre des lignées d'un même génogroupe ou de génogroupes différents. Les substitutions F224Y (VP2) et S93N (VP3) se sont produites respectivement dans les sous-génogroupes C2/C4 et C1/C2. La substitution I249V (VP1) est retrouvée avant l'émergence du sous-génogroupe B2 et durant l'évolution du sous-génogroupe C4. Les substitutions parallèles peuvent conduire à des résidus aminoacides différents dans les lignées, par exemple la position 22 de la protéine VP1 : une substitution Q22R (sous-génogroupe C2) et Q22H (C4). Enfin, les modifications séquentielles de la position 45 de la protéine VP2 est différente entre le génogroupe B (N>T>A>D) et le génogroupe C (N>T>S).

#### **4) La sélection à l'échelle moléculaire parmi les populations d'EV-A71**

La sélection a été analysée à l'échelle moléculaire dans les séquences codant les protéines de la capsid e d'EV-A71, en calculant le rapport des taux de substitutions non synonymes et de substitutions synonymes dN/dS (ou  $\omega$ ) des arbres phylogénétiques estimés par maximum de vraisemblance. Avec les méthodes (SLAC, FEL, REL), les séquences ancestrales des codons sont inférées à chaque site. La fonction de vraisemblance permet de calculer la probabilité des paramètres utilisés pour la modélisation (dont le rapport  $\omega$  pour chaque site) afin de déterminer si les sites sont sélectionnés positivement ou négativement. Les résultats de ces analyses pour l'alignement de la région P1 de l'EV-A71 sont présentés dans le Tableau 13.

Les variations du rapport moyen  $\omega$  pour tous les codons ont été estimées en utilisant la méthode SLAC pour quatre généalogies d'EV-A71 : ce rapport est compris entre 0,023 (EV-A71/C1) et 0,041 (EV-A71/C4). A l'échelle des protéines entières, la proportion de variations non synonymes est significativement très inférieure à celle des changements synonymes dans toutes les populations d'EV-A71. Les changements non synonymes sont soumis à une sélection négative dans les protéines de capsid e, c'est-à-dire à une forte contre-sélection au cours de la transmission interindividuelle des souches.



02

**Figure 47. Phylogramme établi pour le locus 3CD avec 656 séquences d'EV-A**

La phylogénie de l'ensemble des séquences du locus 3CD des 656 souches d'EV-A a été reconstruite avec la méthode du Maximum de Vraisemblance. La fiabilité des nœuds est déterminée par la méthode du SH-test et est indiquée par des étoiles sur les branches. Les branches sont colorées selon les sérotypes. Les flèches indiquent les groupes de souches dont l'origine est potentiellement recombinante.

La proportion de sites sélectionnés négativement a été estimée par les méthodes FEL, IFEL et FUBAR : les pourcentages de sites sur lesquels s'exercent une contre-sélection des changements non-synonymes sont supérieurs à 50% dans toutes les populations d'EV-A71 avec les méthodes FEL et FUBAR. La méthode IFEL rapporte des résultats divergents en estimant des pourcentages <50% dans trois populations.

Ces méthodes permettent aussi d'identifier les sites soumis à une sélection positive [Tableau 13]. Le sous-génogroupe B présente 2 sites : 663 et 710 ; le sous-génogroupe C4 en possède 4 : 416, 663, 710 et 841. La méthode MEME indique qu'un grand nombre de sites sont soumis à une sélection positive épisodique notamment pour les sous-génogroupes B (n=16) et C4 (n=14) par rapport aux sous-génogroupes C1 (n=8) et C2 (n=7). La position 710 est un site épisodique commun à tous les sous-génogroupes d'EV-A71.

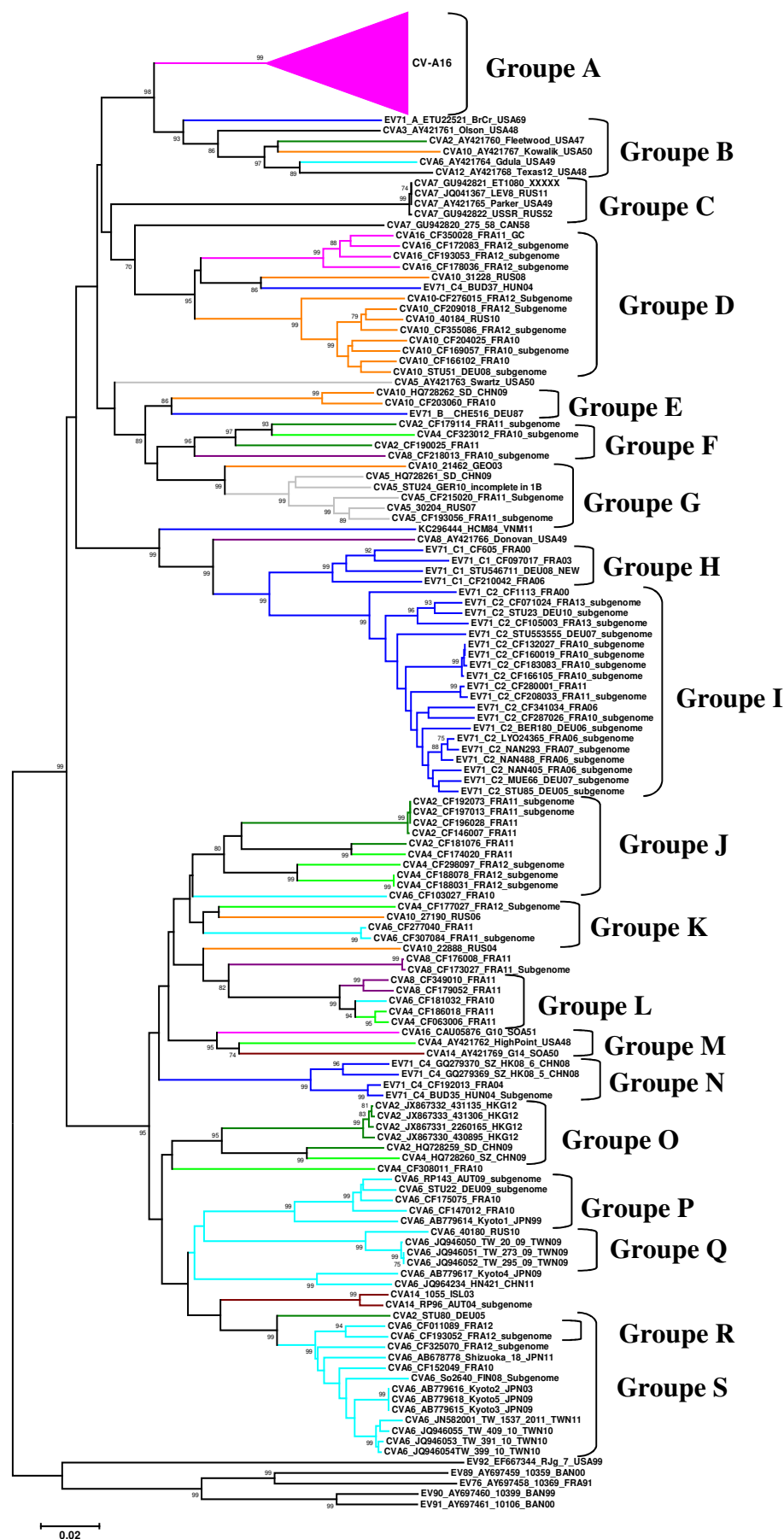
### **III. Recombinaison génétique et diversité des sérotypes de l'espèce EV-A**

#### **1) Recherche des souches d'EV-A d'origine recombinante – Jeux de séquences**

Dans un premier temps, les séquences du locus 3CD d'un échantillon de 394 souches collectées dans des pays européens, échantillonnées parmi plusieurs sérotypes, ont été déterminées pour identifier les souches d'origine recombinante.

#### **A. Première phase de l'étude : analyse phylogénétique exploratoire**

Les séquences de la région génomique non structurale 3CD déterminées chez 394 souches d'EV-A ont été comparées à celles disponibles dans la base de données GenBank (jeu de données constitué en juin 2013, n=686 séquences), par analyse phylogénétique avec la méthode du maximum de vraisemblance [Figure 51]. La phylogénie comprend huit clades fiables ( $p_{SH-test} > 0,7$ ). Contrairement aux clades estimés dans les phylogénies établies pour le locus 1D<sup>VP1</sup>, tous les clades de la phylogénie 3CD ne sont pas monophylétiques. Les clades n° 1, 4, 7 et 8 contiennent des séquences issues de sérotypes différents et sont donc polyphylétiques. Par exemple, le clade n° 7 est constitué de souches de 5 sérotypes différents (CV-A2, CV-A4, CV-A6, CV-A8, EV-A71). Le clade n° 4 est celui qui présente la plus grande diversité de sérotypes (n=8 ; CV-A2, CV-A4, CV-A5, CV-A6, CV-A8, CV-A10, CV-



**Figure 52. Groupes de souches d'origines recombinantes suspectés établis à partir de la phylogénie du locus 3CD des 230 EV-A étudiés.**

La phylogénie de l'ensemble des séquences du locus 3CD des 230 souches d'EV-A a été reconstruite à partir de la méthode du Neighbour Joining. La fiabilité des nœuds est déterminée par la méthode du Bootstrap (1000 réplicats). Les branches sont colorées en fonction des sérotypes.



**Tableau 14.** Sites de recombinaison intertypique estimés parmi les séquences subgénomiques d'EV-A avec sept méthodes d'analyse.

Groupe <sup>a</sup>	N° Recomb	Positions du site de recombinaison		Souche d'origine recombinante	Parent majeur	Parent mineur	Méthodes de détection et valeur P						
		Début	Fin				RDP	GENECO NV	Bootscan	Maxchi	Chimera	SiSscan	3seq
O	1	1	892	CV-A2 HQ728259_CHN09	CV-A2 JX867330_HKG12 81,7 %	CV-A4 HQ728260_CHN09 91,2 %	$8,6 \times 10^{-30}$	$7,9 \times 10^{-24}$	NS	$1,6 \times 10^{-20}$	$2,2 \times 10^{-9}$	NS	$4,1 \times 10^{-44}$
F	2	1799	3905	CV-A2 CF179114_FRA11	CV-A2 CF190025_FRA11 88,3 %	CV-A4 CF323012_FRA10 96,8 %	$3,1 \times 10^{-24}$	$5,5 \times 10^{-19}$	NS	$5,7 \times 10^{-20}$	$2,5 \times 10^{-21}$	NS	$1,3 \times 10^{-10}$
J	3	1902	ND	CV-A2 CF181076_FRA11	CV-A2 CF190025_FRA11 85 %	CV-A4 CF174020_FRA11 96,4 %	$2 \times 10^{-55}$	$1 \times 10^{-57}$	NS	$2,3 \times 10^{-35}$	$2,6 \times 10^{-13}$	$1,2 \times 10^{-50}$	$1,7 \times 10^{-2}$
R	4	2272	3996 <sup>b</sup>	CV-A2 STU80_DEU05	ND	CV-A6 AB77916_JPN03 92,8 %	$1 \times 10^{-37}$	NS	NS	$3,3 \times 10^{-28}$	$1,8 \times 10^{-25}$	$4,2 \times 10^{-21}$	$3,3 \times 10^{-57}$
L	5	1384	3996 <sup>b</sup>	CV-A4 CF186018_FRA11	CV-A4 CF323012_FRA10 94,1 %	CV-A8 CF179052_FRA11 96,4 %	$1,2 \times 10^{-84}$	$5 \times 10^{-79}$	NS	$4,1 \times 10^{-41}$	$1,4 \times 10^{-42}$	$6,7 \times 10^{-50}$	$6 \times 10^{-160}$
L	6	1384	3996 <sup>b</sup>	CV-A4 CF063006_FRA11	CV-A4 CF323012_FRA10 94,1 %	CV-A8 CF179052_FRA11 96,4 %	$1,2 \times 10^{-84}$	$5 \times 10^{-79}$	NS	$4,1 \times 10^{-41}$	$1,4 \times 10^{-42}$	$6,7 \times 10^{-50}$	$6 \times 10^{-160}$
J	7	1552	2830	CV-A4 CF177027_FRA12	CV-A10 27190_RUS06 77,4 %	CV-A6 CF277040_FRA11 92,7 %	$4,6 \times 10^{-15}$	$1,2 \times 10^{-6}$	NS	$8,3 \times 10^{-14}$	$6,9 \times 10^{-10}$	NS	$4,8 \times 10^{-2}$
K	8	1876	4017	CV-A4 CF174020_FRA11	ND	CV-A2 CF181076_FRA11 96,2 %	$5,9 \times 10^{-37}$	$3,9 \times 10^{-27}$	NS	$2,2 \times 10^{-27}$	$6,9 \times 10^{-31}$	NS	$4,5 \times 10^{-47}$
S	9	401	921	CV-A6 JQ946053_TWN10	CV-A6 JN582001_TWN11 99,1 %	ND	$4,4 \times 10^{-29}$	$2,9 \times 10^{-27}$	NS	$5,5 \times 10^{-10}$	$5 \times 10^{-10}$	$2,6 \times 10^{-11}$	$4,2 \times 10^{-30}$
K	10	1554	1879	CV-A6 CF277040_FRA11	CV-A6 CF325070_FRA12 82,9 %	CV-A4 CF177027_FRA12 95,1 %	$2,3 \times 10^{-30}$	$4,2 \times 10^{-19}$	NS	$1,7 \times 10^{-10}$	$2 \times 10^{-11}$	$4,3 \times 10^{-9}$	$7,3 \times 10^{-25}$
K	11	1554	1879	CV-A6 CF307084_FRA11	CV-A6 CF325070_FRA12 82,9 %	CV-A4 CF177027_FRA12 95,1 %	$2,3 \times 10^{-30}$	$4,2 \times 10^{-19}$	NS	$1,7 \times 10^{-10}$	$2 \times 10^{-11}$	$4,3 \times 10^{-9}$	$7,3 \times 10^{-25}$
	12	2271	3996 <sup>b</sup>	CV-A6 CF103027_FRA10	CV-A6 CF175075_FRA10 93,1 %	ND	$1,6 \times 10^{-2}$	NS	NS	$3,8 \times 10^{-25}$	$8,6 \times 10^{-14}$	$2 \times 10^{-4}$	$3,5 \times 10^{-4}$

Groupe	N° Recomb	Positions du site de recombinaison		Souche d'origine recombinante	Parent majeur	Parent mineur	Méthodes de détection et valeur P						
		Début	Fin				RDP	GENECO NV	Bootscan	Maxchi	Chimera	SiSscan	3seq
L	13	2844	3996 <sup>b</sup>	CV-A6 CF181032_FRA10	CV-A6 CF307084_FRA11 87,7 %	CV-A4 CF063006_FRA11 98,1 %	6,5x10 <sup>-49</sup>	1,4x10 <sup>-31</sup>	NS	5,3x10 <sup>-26</sup>	5,2x10 <sup>-27</sup>	NS	1,2x10 <sup>-45</sup>
F	14	1018	3996 <sup>b</sup>	CV-A8 CF218013_FRA10	CV-A8 Donovan_USA49 81,9 %	ND	1,1x10 <sup>-27</sup>	4,6x10 <sup>-17</sup>	8,8x10 <sup>-10</sup>	3,3x10 <sup>-15</sup>	1,1x10 <sup>-12</sup>	3x10 <sup>-20</sup>	1,9x10 <sup>-5</sup>
	15	1082	3996 <sup>b</sup>	CV-A8 CF176008_FRA11	CV-A8 CF218013_FRA10 90 %	ND	7,6x10 <sup>-48</sup>	2,3x10 <sup>-41</sup>	NS	1,8x10 <sup>-17</sup>	7,7x10 <sup>-25</sup>	3,3x10 <sup>-39</sup>	7,4x10 <sup>-36</sup>
	16	1082	956	CV-A8 CF173027_FRA11	CV-A8 CF218013_FRA10 90 %	ND	7,6x10 <sup>-48</sup>	2,3x10 <sup>-41</sup>	NS	1,8x10 <sup>-17</sup>	7,7x10 <sup>-25</sup>	3,3x10 <sup>-39</sup>	7,4x10 <sup>-36</sup>
D	17	2307	3708	CV-A10 31228_RUS08	CV-A10 CF169057_FRA10 92,1 %	CV-A16 CF350028_FRA11 91,6 %	4,7x10 <sup>-27</sup>	NS	4,3x10 <sup>-26</sup>	3,1x10 <sup>-14</sup>	2,1x10 <sup>-3</sup>	NS	1,5x10 <sup>-22</sup>
G	18	3537	3996 <sup>b</sup>	CV-A10 21462_GEO03	ND	CV-A5 30204_RUS07 98,9 %	1,6x10 <sup>-75</sup>	9,2x10 <sup>-72</sup>	NS	1,5x10 <sup>-18</sup>	5,1x10 <sup>-20</sup>	1,6x10 <sup>-21</sup>	2,9x10 <sup>-68</sup>
A	19	1754	1886	CV-A16 40188_RUS04	CV-A16 40187_RUS04 99,3 %	ND	3,5x10 <sup>-17</sup>	1,1x10 <sup>-14</sup>	NS	6,4x10 <sup>-6</sup>	2,4x10 <sup>-6</sup>	5,9x10 <sup>-7</sup>	1,5x10 <sup>-12</sup>
A	20	2920	3996 <sup>b</sup>	CV-A16 KF193630_CHN10	CV-A16 KF193629_CHN10 96 %	ND	1,9x10 <sup>-24</sup>	7,4x10 <sup>-21</sup>	NS	1x10 <sup>-16</sup>	7,8x10 <sup>-19</sup>	1,7x10 <sup>-21</sup>	NS
102	21	3442	3996 <sup>b</sup>	EV-A71/C4 BUD37_HUN04	EV-A71/C4 CF192013_FRA04 99,4 %	CV-A10 31228_RUS08 93,7 %	8,9x10 <sup>-70</sup>	5,8x10 <sup>-39</sup>	3,5x10 <sup>-63</sup>	1,8x10 <sup>-21</sup>	7,3x10 <sup>-21</sup>	1,2x10 <sup>-17</sup>	1,6x10 <sup>-47</sup>

<sup>a</sup> Les groupes étudiés sont représentés sur l'arbre phylogénétique de la Figure 52.

<sup>b</sup> La position 3996 indique la fin du segment sugénomique étudié.

Abréviations : ND, non déterminé ; NS, non significatif (P<10<sup>-5</sup>).

**Tableau 15.** Sites de recombinaison intratypique estimés parmi les séquences subgénomiques d'EV-A sept méthodes d'analyse.

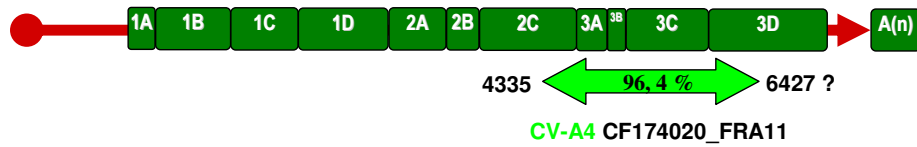
Groupe <sup>a</sup>	N° Recomb	Positions du site de recombinaison		Souche d'origine recombinante	Parent mineur	Parent majeur	Valeur P pour chaque méthodes d'analyse						
		Début	Fin				RDP	GENECO NV	Bootscan	Maxchi	Chimera	SiSscan	3seq
J	22	485	1029	CV-A4 CF188078_FRA12	CV-A4 CF188031_FRA12 98,9 %	CV-A4 CF298097_FRA12 95,6 %	$3,1 \times 10^{-10}$	NS	NS	$1,9 \times 10^{-9}$	$5,1 \times 10^{-10}$	NS	$4,4 \times 10^{-5}$
S	23	1564	3996 <sup>b</sup>	CV-A6 AB779616_JPN03	CV-A6 AB779614_JPN09 96,2 %	CV-A6 AB776918_JPN09 100 %	$1,2 \times 10^{-17}$	NS	NS	$2 \times 10^{-20}$	$4,8 \times 10^{-15}$	$1,3 \times 10^{-5}$	$1,1 \times 10^{-39}$
D	24	923	1865	CV-A10 CF169057_FRA10	CV-A10 40184_RUS10 96,8 %	CV-A10 CF276015_FRA12 99,7 %	$1,4 \times 10^{-22}$	$3,6 \times 10^{-20}$	$3,2 \times 10^{-23}$	$7,1 \times 10^{-13}$	$3,3 \times 10^{-12}$	NS	$3,3 \times 10^{-17}$
A	25	824	3318	CV-A16 STU54_DEU08	CV-A16 KC117318_CHN09 95,8 %	CV-A16 BUD25_HUN08 98,1 %	$1,5 \times 10^{-12}$	$2,8 \times 10^{-10}$	NS	$7,4 \times 10^{-14}$	$3,2 \times 10^{-13}$	$4,4 \times 10^{-26}$	$1,1 \times 10^{-19}$
D	26	2319	3201	CV-A16 CF193053_FRA12	CV-A16 CF350028_FRA11 97,3 %	ND	$9,5 \times 10^{-62}$	$1,4 \times 10^{-41}$	$1,9 \times 10^{-55}$	$1,8 \times 10^{-55}$	$6,6 \times 10^{-62}$	NS	$9,5 \times 10^{-62}$

<sup>a</sup> Les groupes étudiés sont représentés sur l'arbre phylogénétique de la **Figure 52**.

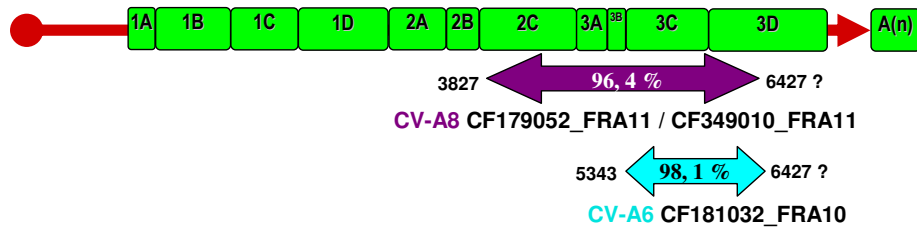
<sup>b</sup> La position 3996 indique la fin du segment sugénomique étudié.

Abréviations : ND, non déterminé ; NS, non significatif ( $P < 10^{-5}$ ).

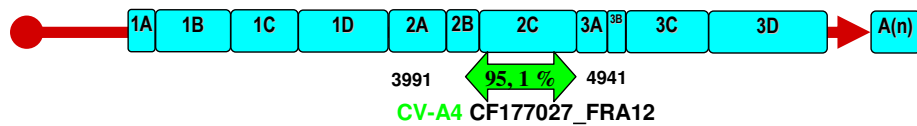
N°3 / 8 : CV-A2 CF181076\_FRA11



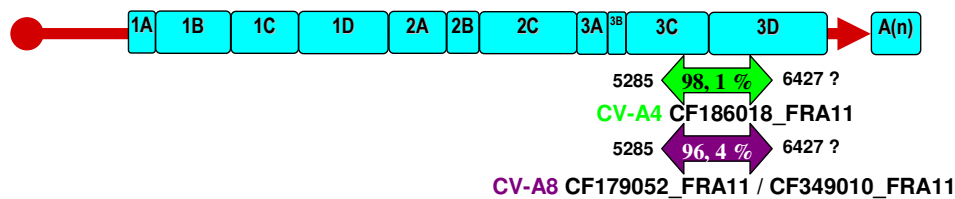
N°5 / 6 : CV-A4 CF186018\_FRA11 / CF063006\_FRA11



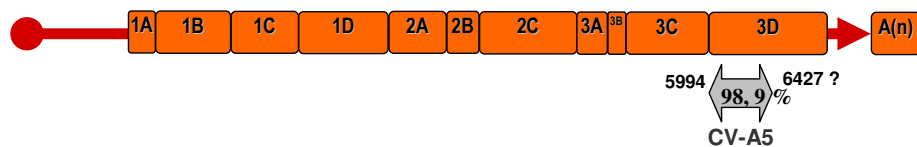
N°10 / 11 : CV-A6 CF277040\_FRA11 / CF307084\_FRA11



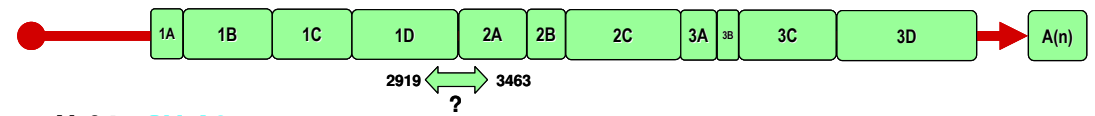
N°13 : CV-A6 CF181032\_FRA11



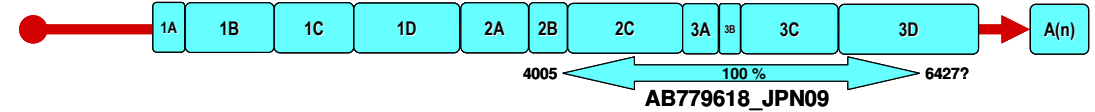
N°18 : CV-A10 21462\_GEO03



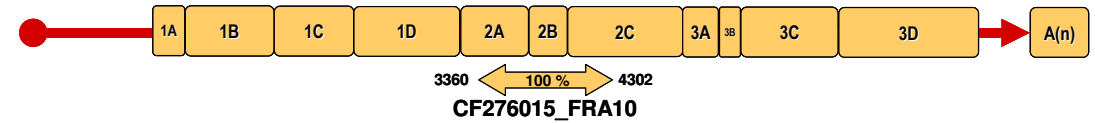
N°23 : CV-A4 CF188078\_FRA12



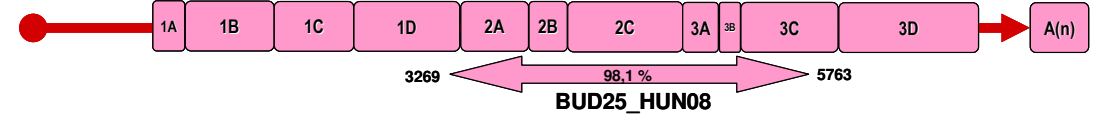
N°24 : CV-A6 AB779616\_JPN03



N°25 : CV-A10 CF169057\_FRA10



N°26 : CV-A16 STU54\_DEU08



N°26 : CV-A16 CF193053\_FRA12

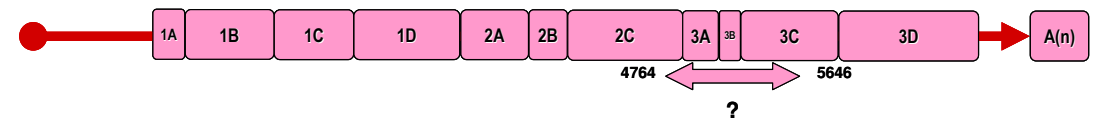


Figure 53. Schéma récapitulatif des évènements de recombinaisons intertypiques (A) et intratypiques (B) mis en évidence pour les EV-A

A16, EV-A71). Les quatre clades polyphylétiques comprennent une centaine de souches parmi lesquelles certaines ont potentiellement une origine recombinante.

## **B. Séquençage génomique et constitution des jeux de données**

A partir de l'analyse précédente, 113 souches ont été sélectionnées pour poursuivre cette étude. Le génome complet a été déterminé pour 38 souches et les séquences subgénomiques de 4000 nucléotides (depuis le locus 1D<sup>VP1</sup> jusqu'aux gènes 3CD) ont été déterminées pour 75 souches. Deux méthodes de PCR ont été employées au cas par cas : PCR en deux parties (5 CV-A2, 4 CV-A4, 1 CV-A5, 7 CV-A6, 3 CV-A8, 3 CV-A10, 1 CV-A14, 4 CV-A16 et 7 EV-A71) ou PCR simple subgénomique entre les loci 1D<sup>VP1</sup> et 3CD (3 CV-A2, 5 CV-A4, 3 CV-A5, 7 CV-A6, 2 CV-A8, 12 CV-A10, 1 CV-A14, 21 CV-A16 et 21 EV-A71). La principale difficulté dans la réalisation de ce travail de séquençage fut la variabilité génétique importante des souches qui limitait l'utilisation d'amorces pour de nombreuses souches du même sérotype. A ces 113 séquences subgénomiques d'EV-A, ont été ajoutées les séquences disponibles de génomes complets d'EV-A pour former un alignement de 230 séquences subgénomiques (quelques séquences d'EV-A71 ont été sélectionnées pour chaque sous-génogroupe pour ne pas biaiser l'échantillon en faveur de ce sérotype). Une comparaison des topologies obtenues par « Neighbour Joining » à partir des loci 1D et 3CD des 230 séquences d'EV-A a été effectuée pour contrôler l'échantillon. Les incongruences phylogénétiques entre les deux arbres sont identiques à celles mises en évidence dans l'analyse précédente. Les groupes de recombinaison ont été désignés par des lettres [Figure 52]. Une recherche de signature moléculaire de recombinaison a ensuite été réalisée afin de prouver leur origine recombinante.

## **2) Analyse des signatures d'événements de recombinaison génétique**

### **A. Analyse des séquences subgénomiques**

L'analyse du jeu de données de 230 séquences subgénomiques a révélé la présence de 26 sites de recombinaison fiables entre les loci 1D<sup>VP1</sup> et 3CD. La fiabilité de ces sites repose sur l'obtention d'une valeur de probabilité proche de zéro ( $P < 10^{-5}$ ) par au moins 3 des 7 méthodes d'analyse utilisées (RDP, GENECONV, Bootscan, MaxChi, Chimera, Siscan et 3SEQ). Les sites peuvent être classés en deux catégories d'événements de recombinaison.

La première catégorie comprend 21 sites d'événements de recombinaison intertypique (Tableau 14) impliquant des séquences des 7 sérotypes suivants : CV-A2, CV-A4, CV-A5, CV-A6, CV-A8, CV-A10 et CV-A16 [Figure 53]. Les sites de recombinaison ont été estimés

**Tableau 16. Détection des sites de recombinaison intratypique avec 7 méthodes d'analyse parmi les séquences ORF de coxsackievirus A16<sup>a</sup>.**

N° du site de recombinaison	Positions du site de recombinaison		Séquence d'origine recombinante	Séquence parentale majeure	Séquence parentale mineure	Valeur P pour chaque méthode d'analyse						
	Début	Fin				RDP	GENECONV	Bootscan	Maxchi	Chimera	SiScan	3seq
27	641	1735	CV-A16 JX986742_CHN11	CV-A16 KC755231_CHN10 97,8 %	CV-A16 JX986740_CHN11 99,7 %	7,2x10 <sup>-21</sup>	9,7x10 <sup>-21</sup>	NS	4,5x10 <sup>-10</sup>	2,5x10 <sup>-6</sup>	1,4x10 <sup>-14</sup>	1,2x10 <sup>-15</sup>
28	4605	6556 <sup>b</sup>	CV-A16 KF193630_CHN10	CV-A16 KF055238_CHN10 98,7 %	CV-A16 KC755230_CHN10 97 %	1,8x10 <sup>-40</sup>	6,8x10 <sup>-22</sup>	NS	876x10 <sup>-25</sup>	5,1x10 <sup>-24</sup>	1,3x10 <sup>-20</sup>	NS
29	4685	5899	CV-A16 KF193629_CHN10	CV-A16 JX068832_CHN06 98,7 %	ND	3x10 <sup>-13</sup>	NS	NS	3,2x10 <sup>-13</sup>	3,2x10 <sup>-12</sup>	7,4x10 <sup>-9</sup>	4,8x10 <sup>-18</sup>
30	5733	5928	CV-A16 KF193627_CHN09	CV-A16 KC755231_CHN10 98,3 %	CV-A16 G10_SOA51 100 %	2,4x10 <sup>-51</sup>	6,1x10 <sup>-50</sup>	NS	3x10 <sup>-10</sup>	2,2x10 <sup>-10</sup>	1,8x10 <sup>-11</sup>	NS

<sup>a</sup> Les sites des recombinaison décrits dans le tableau ont été estimés parmi les séquences du groupe phylogénétique A (**Figure 52**).

<sup>b</sup> La position 6556 indique la fin du segment génétique (ORF) étudié.

Abréviations : ND, non déterminé ; NS, non significatif (P<10<sup>-5</sup>).

**Table 17 :** Evénements de recombinaison potentiels détectés par RDP4 au sein de l’alignement de l’ORF des EV-A71.

Groupe	N° Recomb	Positions du site de recombinaison		Souche d’origine recombinante	Parent majeur	Parent mineur	Méthodes de détection et valeur P						
		Début	Fin				RDP	GENECO NV	Bootscan	Maxchi	Chimera	SiSscan	3seq
B0	31	3535	6579 <sup>a</sup>	EV-A71/B0 AB575911_NLD66	EV-A71/B1 AB575913_NLD71 94 %	ND	3x10 <sup>-20</sup>	1,8x10 <sup>-13</sup>	1,1x10 <sup>-25</sup>	8,6x10 <sup>-10</sup>	1,6x10 <sup>-12</sup>	4,6x10 <sup>-27</sup>	4,9x10 <sup>-14</sup>
B0	32	3535	6579 <sup>a</sup>	EV-A71/B0 AB575912_NLD66	EV-A71/B1 AB575913_NLD71 94 %	ND	3x10 <sup>-20</sup>	1,8x10 <sup>-13</sup>	1,1x10 <sup>-25</sup>	8,6x10 <sup>-10</sup>	1,6x10 <sup>-12</sup>	4,6x10 <sup>-27</sup>	4,9x10 <sup>-14</sup>
B3	33	3644	6579 <sup>a</sup>	EV-A71/B3 EU364841_AUS99	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	34	3644	6579 <sup>a</sup>	EV-A71/B3 AM396586_MYS97	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	35	3644	6579 <sup>a</sup>	EV-A71/B3 AM396588_MYS97	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	36	3644	6579 <sup>a</sup>	EV-A71/B3 DQ341354_SGP98	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	37	3644	6579 <sup>a</sup>	EV-A71/B3 DQ341367_MYS97	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	38	3644	6579 <sup>a</sup>	EV-A71/B3 DQ341368_MYS97	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B3	39	3644	6579 <sup>a</sup>	EV-A71/B3 AB469182_MYS97	EV-A71/B3 AJ586873_MYS97 95,8 %	ND	1,2x10 <sup>-31</sup>	1,5x10 <sup>-10</sup>	1,1x10 <sup>-12</sup>	7,6x10 <sup>-19</sup>	8,9x10 <sup>-20</sup>	1,4x10 <sup>-22</sup>	1,7x10 <sup>-39</sup>
B5	40	1821	2374	EV-A71/B5 JF738001_THA09	EV-A71/B5 FJ461781_SGP08 99,1 %	ND	1,3x10 <sup>-35</sup>	9x10 <sup>-7</sup>	NS	1,7x10 <sup>-11</sup>	2x10 <sup>-10</sup>	4,5x10 <sup>-5</sup>	4,2x10 <sup>-9</sup>
C1	41	4708	6579 <sup>a</sup>	EV-A71/C1 EF063152_TWN05	EV-A71/C1 AB575936_NLD91 90,8 %	ND	NS	NS	6,3x10 <sup>-4</sup>	1,3x10 <sup>-7</sup>	1,3x10 <sup>-10</sup>	6,1x10 <sup>-17</sup>	NS
C2	42	1810	2390	EV-A71/C2 AF119795_TWN98	EV-A71/C2 AF119796_TWN98 98,5 %	ND	9,9x10 <sup>-28</sup>	3x10 <sup>-12</sup>	2,6x10 <sup>-19</sup>	3,7x10 <sup>-14</sup>	3,5x10 <sup>-12</sup>	2,7x10 <sup>-10</sup>	3,4x10 <sup>-17</sup>
C2 (like)	43	3265	6579 <sup>a</sup>	EV-A71/C2 JQ280307_TWN08	ND	ND	7,3x10 <sup>-6</sup>	NS	1,6x10 <sup>-2</sup>	1,3x10 <sup>-24</sup>	2,8x10 <sup>-15</sup>	2x10 <sup>-20</sup>	9,3x10 <sup>-9</sup>
C2 (like)	44	3265	6579 <sup>a</sup>	EV-A71/C2 HM622391_TWN08	ND	ND	7,3x10 <sup>-6</sup>	NS	1,6x10 <sup>-2</sup>	1,3x10 <sup>-24</sup>	2,8x10 <sup>-15</sup>	2x10 <sup>-20</sup>	9,3x10 <sup>-9</sup>

Groupe	N° Recomb	Positions du site de recombinaison		Souche d'origine recombinante	Parent majeur	Parent mineur	Méthodes de détection et valeur P						
		Début	Fin				RDP	GENECO NV	Bootscan	Maxchi	Chimera	SiScan	3seq
C2 (like)	45	3265	6579 <sup>a</sup>	EV-A71/C2 HM622392_TWN08	ND	ND	7,3x10 <sup>-6</sup>	NS	1,6x10 <sup>-2</sup>	1,3x10 <sup>-24</sup>	2,8x10 <sup>-15</sup>	2x10 <sup>-20</sup>	9,3x10 <sup>-9</sup>
C4	46	766	1865	EV-A71/C4 JX244187_CHN10	EV-A71/C4 JX244188_CHN10 99,3 %	EV-A71/C4 JX244182_CHN10	7x10 <sup>-17</sup>	3,3x10 <sup>-15</sup>	NS	3,3x10 <sup>-8</sup>	3,6x10 <sup>-8</sup>	1,5x10 <sup>-8</sup>	1,7x10 <sup>-20</sup>
C4	47	854 ?	1824	EV-A71/C4 KC954664_CHN09	ND	EV-A71/C4 KC954662_CHN07 100 %	3,5x10 <sup>-42</sup>	1,2x10 <sup>-53</sup>	NS	2x10 <sup>-14</sup>	7,5x10 <sup>-8</sup>	2,4x10 <sup>-23</sup>	4,6x10 <sup>-30</sup>
C4	48	2520	3588	EV-A71/C4 FJ360546_CHN08	EV-A71/C4 JX678879_CHN08 99,1 %	EV-A71/C4 FJ360545_CHN08 99,5 %	9,6x10 <sup>-12</sup>	1,4x10 <sup>-9</sup>	NS	6x10 <sup>-5</sup>	1x10 <sup>-5</sup>	3x10 <sup>-7</sup>	1,6x10 <sup>-14</sup>
C4	49	3095	3455	EV-A71/C4 AF302996_CHN98	EV-A71/C4 HQ694983_CHN08 97,4 %	ND	1,1x10 <sup>-9</sup>	1,3x10 <sup>-7</sup>	NS	7,2x10 <sup>-5</sup>	8,3x10 <sup>-5</sup>	NS	NS
C4	50	3380	6579 <sup>a</sup>	EV-A71/C4 JF799986_CHN09	EV-A71/C4 HQ129932_CHN06 98,6 %	ND	3,6x10 <sup>-29</sup>	4,2x10 <sup>-28</sup>	NS	1,1x10 <sup>-16</sup>	2,4x10 <sup>-19</sup>	9,3x10 <sup>-34</sup>	4x10 <sup>-42</sup>
C4	51	3699	4872	EV-A71/C4 HM807310_TWN05	EV-A71/C4 GQ231938_TWN05 98,7 %	ND	4,6x10 <sup>-44</sup>	7,4x10 <sup>-42</sup>	NS	7,1x10 <sup>-26</sup>	1,4x10 <sup>-27</sup>	1,6x10 <sup>-25</sup>	2,2x10 <sup>-50</sup>
C4	52	5072	6579 <sup>a</sup>	EV-A71/C4 FJ606450_CHN08	EV-A71/C4 HQ129932_CHN06 97,5 %	ND	2,8x10 <sup>-6</sup>	NS	NS	2,2x10 <sup>-13</sup>	4,2x10 <sup>-5</sup>	1,8x10 <sup>-9</sup>	3x10 <sup>-18</sup>
C4	53	5098	6579 <sup>a</sup>	EV-A71/C4 FJ607337_CHN08	EV-A71/C4 HQ129932_CHN06 97,1 %	ND	8x10 <sup>-9</sup>	3,8x10 <sup>-12</sup>	NS	8,4x10 <sup>-11</sup>	1,2x10 <sup>-11</sup>	1,7x10 <sup>-13</sup>	1,7x10 <sup>-18</sup>
C4	54	5134	6579 <sup>a</sup>	EV-A71/C4 BUD37_HUN04	EV-A71/C4 CF192013_FRA04 99,4 %	ND	5,2x10 <sup>-32</sup>	5,8x10 <sup>-30</sup>	NS	6,2x10 <sup>-21</sup>	2,1x10 <sup>-20</sup>	5,1x10 <sup>-23</sup>	1,5x10 <sup>-69</sup>
C4	55	5370	6579 <sup>a</sup>	EV-A71/C4 HQ423142_CHN09	EV-A71/C4 FJ194964_CHN08 96,4 %	ND	5,6x10 <sup>-11</sup>	2,7x10 <sup>-5</sup>	NS	2,3x10 <sup>-7</sup>	3,9x10 <sup>-10</sup>	6,3x10 <sup>-16</sup>	NS
C4	56	5370	6579 <sup>a</sup>	EV-A71/C4 HQ423143_CHN09	EV-A71/C4 FJ194964_CHN08 96,4 %	ND	5,6x10 <sup>-11</sup>	2,7x10 <sup>-5</sup>	NS	2,3x10 <sup>-7</sup>	3,9x10 <sup>-10</sup>	6,3x10 <sup>-16</sup>	NS
C4	57	5370	6579 <sup>a</sup>	EV-A71/C4 KJ784495_CHN11	EV-A71/C4 FJ194964_CHN08 95,6 %	ND	5,6x10 <sup>-11</sup>	2,7x10 <sup>-5</sup>	NS	2,3x10 <sup>-7</sup>	3,9x10 <sup>-10</sup>	6,3x10 <sup>-16</sup>	NS

<sup>a</sup> La position 6579 indique la fin du segment génétique (ORF) étudié.  
Abréviations : ND, non déterminé ; NS, non significatif (P<10<sup>-5</sup>).



majoritairement au niveau du gène de la protéine non structurale 2C (n=10/21). Un seul site de recombinaison est localisé dans le locus 1D<sup>VP1</sup> (site n° 9). Lorsque la similitude nucléotidique entre deux séquences est < 95%, les méthodes d'analyse ne permettent pas d'identifier les génomes parentaux (séquences parentales donatrices) à partir desquels se sont produits les événements de recombinaison. Dans notre échantillon, les génomes parentaux n'ont pas été identifiés pour les recombinaisons n° 4, 7, 9, 12, 14, 15, 16 et 19.

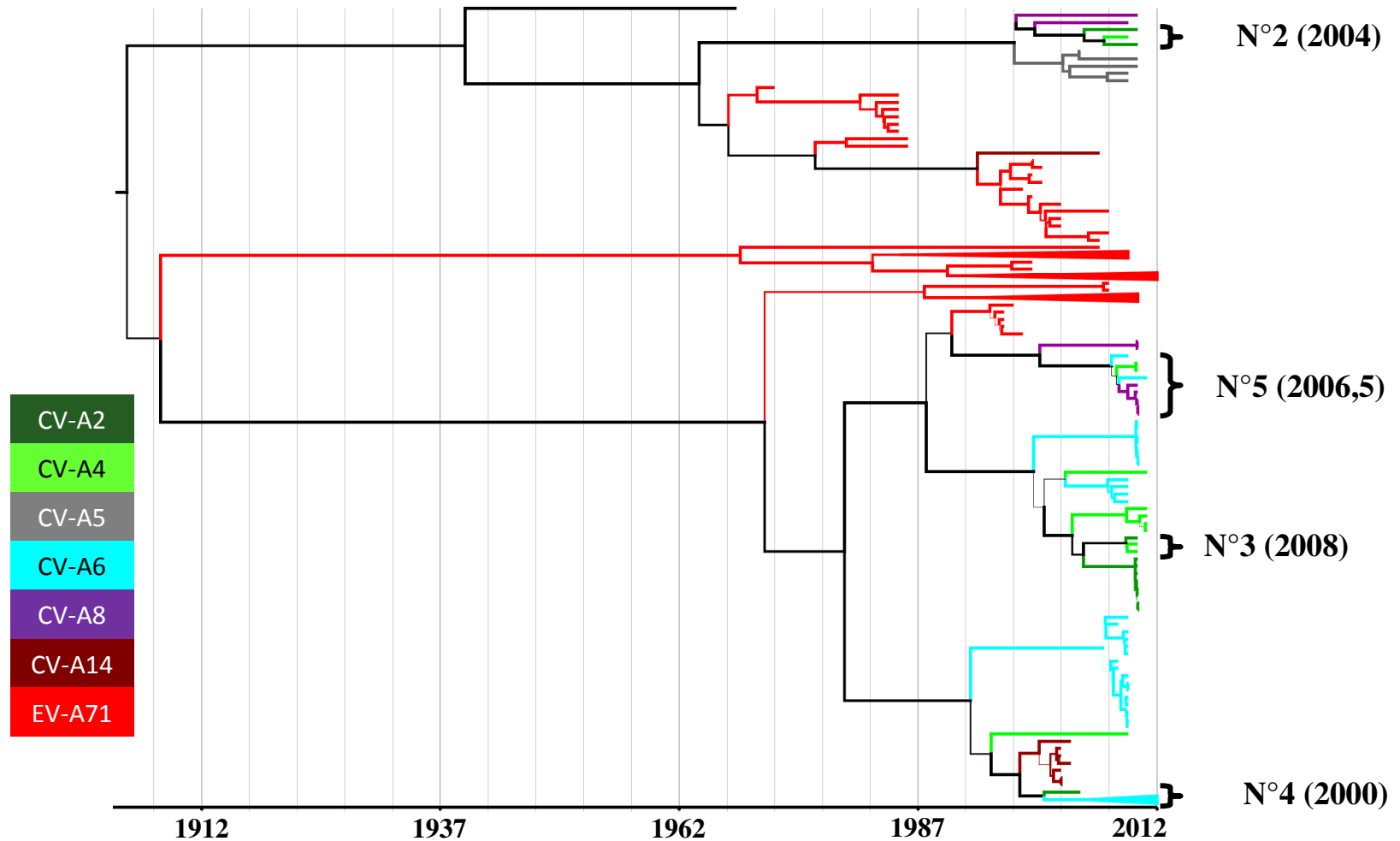
La seconde catégorie rassemble cinq sites d'évènements de recombinaison intratypique estimés entre des séquences de quatre sérotypes : CV-A4, CV-A6, CV-A10 et CV-A16 (Tableau 15). Les sites de recombinaison intratypique sont localisés dans le locus 1D<sup>VP1</sup> (n=2) et dans chacun des gènes non structuraux 2A, 2B et 2C. Les séquences parentales donatrices des recombinaisons n° 23 et 26 n'ont pas pu être identifiées.

### **B. Analyse des séquences de l'ORF du CV-A16**

Pour étudier la recombinaison génétique parmi les CV-A16, un jeu de données rassemblant 71 séquences de l'ORF a été constitué en utilisant les séquences déterminées dans cette étude (n = 3) et celles disponibles dans la base de données GenBank (n = 68). L'analyse de cet alignement a permis d'estimer la présence de 4 sites de recombinaison ( $P < 10^{-5}$ ) avec au moins 5 des 7 méthodes d'analyse implémentées dans le programme RDP [Tableau S4]. Trois sites (n° 27, 28 et 30) représentent des événements de recombinaison intratypique car les deux séquences donatrices étaient présentes dans le jeu de données, tandis que pour le site n° 29, seule l'une d'entre elles a pu être identifiée. Les sites de recombinaison ont été localisés dans la région génomique P3 (3A à 3D) à l'exception du site n° 27 qui est situé dans la région P1 (locus 1B). L'évènement de recombinaison intratypique n° 30 a la particularité d'avoir la souche prototype CV-A16 comme souche parentale donatrice d'un fragment de 195 nucléotides avec une similitude de 100%), ce qui suggère une contamination ou une erreur de séquençage.

### **C. Analyse des séquences de l'ORF de l'EV-A71**

Une étude similaire à la précédente a été effectuée pour l'EV-A71 avec un jeu de données rassemblant 541 séquences de l'ORF (séquences déterminées dans cette étude, n = 10 ; séquences issues de la base de données GenBank, n = 531). L'analyse des séquences ORF d'EV-A71 a permis d'estimer 27 sites de recombinaison au seuil statistique  $P < 10^{-5}$ , avec au moins 5 des 7 méthodes du programme RDP4 [Tableau S5]. Les sites n° 46, 47, 48 constituent des événements de recombinaison intratypique car deux séquences parentales



**Figure 54. Chronogramme des EV-A inféré pour le locus 3CD, permettant la datation des évènements de recombinaison**

Le chronogramme a été inféré par une approche bayésienne et un modèle d’horloge moléculaire assouplie, en utilisant un jeu de données pour les sérotypes indiqués. Les noms des taxons ne sont pas indiqués. Les épaisseurs des branches sont proportionnelles aux probabilités postérieures. Les branches sont colorées en fonction des sérotypes. Les accolades indiquent les groupes de recombinaison. Les dates des évènements de recombinaison sont indiquées entre parenthèses.

donatrices ont été détectées parmi les séquences d'EV-A71 (sous-génogroupe C4). Pour les 24 autres sites estimés, l'une des séquences parentales donatrices ne peut pas être identifiée dans le jeu de données utilisé [Tableau S5]. Certains événements de recombinaison se retrouvent au sein de toutes les souches d'un même sous-génogroupe tels que les événements n°31/32 et 33-39 qui concernent respectivement les sous-génogroupes EV-A71/B0 et EV-A71/B3. Les sites des événements de recombinaison intratypiques sont localisés dans la région P1 (n=3) alors que les autres sont principalement dans P2 (n=15) et P3 (n=7), les régions codant les protéines non structurales.

#### **D. Analyse croisée des séquences de l'ORF du CV-A16 et de l'EV-A71**

Les études épidémiologiques concernant la maladie PMB rapportent fréquemment la co-circulation des souches des sérotypes CV-A16 et EV-A71 dans les pays asiatiques [Li *et al*, 2005]. La co-circulation peut favoriser les co-infections et la recombinaison entre ces deux entérovirus, lors des périodes épidémiques par exemple. Nous avons donc recherché la présence de sites de recombinaison par une analyse croisée des jeux de séquences de ces deux sérotypes. Les résultats indiquent l'absence de signal d'évènement de recombinaison intertypique entre l'EV-A71 et le CV-A16 détectables par les méthodes de RDP.

### **3) Datation des événements de recombinaison par analyses bayésiennes**

Les événements de recombinaison peuvent être datés par analyse phylogénétique pour établir la chronologie de leur survenue dans les populations virales en circulation et étudier les liens avec l'épidémiologie (épidémies ou circulation sporadique des populations virales parentales). Les dates de survenue des événements de recombinaison intertypiques ont été estimées par une analyse bayésienne du jeu de données 3CD. Le chronogramme obtenu a permis d'estimer la date des événements de recombinaison n° 2, 3, 4 et 5 [Figure 54]. Les événements de recombinaison n° 2, 3 et 5 qui impliquent des souches isolées en France récemment, se sont produits respectivement en 2004 [2000 ; 2007], 2008 [2006 ; 2009] et 2006,5 [2005 ; 2008] soit 3 à 6 ans avant le premier isolement d'une souche de ces groupes.

La topologie phylogénétique du groupe n° 5 suggère la survenue d'une chaîne de recombinaison entre les souches des sérotypes CV-A4, CV-A6 et CV-A8 [Figure 54]. L'évènement de recombinaison n° 4 qui s'est produit entre une souche de CV-A2 et un cluster de souches de CV-A6 a eu lieu environ en 2000 [1996 ; 2003]. Les autres événements de

recombinaison n'ont pas pu être datés par une approche bayésienne en raison de nœud non fiables ( $pp < 0.90$ ).

#### **IV. Etude du pouvoir pathogène des EV-A avec un modèle cellulaire de barrière endothéliale microvasculaire de derme**

Une lignée cellulaire d'endothélium microvasculaire de derme appelée « HMEC-1 », obtenue auprès du Centre de Contrôle des Maladies Infectieuses (CDC) d'Atlanta (USA), a été testée dans le but de constituer un modèle d'étude des mécanismes physiopathologiques des infections à EV-A. Cette lignée a été obtenue par immortalisation d'une culture primaire de cellules endothéliales à l'aide du virus SV40. Dans la littérature, il s'agit de la principale lignée de cellules endothéliales de derme utilisée dans des domaines scientifiques tels que la physiologie de l'endothélium, l'immunologie (interactions cellules immunitaire / barrière endothéliale) l'étude des dysfonctionnements endothéliaux associés à des pathologies (diabète, ...), l'étude de l'angiogenèse associé à des processus tumoraux, l'étude de la virémie de pathogènes, etc. Le rationnel de cette étude repose sur le fait que la virémie associée à une infection par un virus donnant des éruptions cutanées de maladie PMB est due probablement au franchissement des barrières endothéliales microvasculaires du derme.

##### **1) Mise au point des conditions de culture de la lignée HMEC-1**

La lignée cellulaire « HMEC-1 » (réceptionnée au passage 11) a été mise en culture avec un milieu complet dédié à la culture de cellules endothéliales et en présence de collagène-I. Ces cellules se caractérisent par une morphologie de petite taille et par un tapis cellulaire en forme de « pavés ». Les passages réalisés avec cette lignée (P12 à P25) suggèrent que le temps de doublement est d'environ 24 heures.

Une étude de la cinétique de croissance cellulaire de la lignée HMEC-1 a été réalisée dans des microplaques 96 puits pré-enduites de collagène I. Après avoirensemencé chaque puits avec 10 000 cellules, une numération a été effectuée au bout de 24, 48 et 72 heures de culture. La cinétique suggère une croissance de type exponentiel après 48 heures de culture (données non montrées).

**Tableau 18.** Caractéristiques des infections des cellules HMEC-1 pour 42 souches d'entérovirus

#	Souche	Serotype	Pays	Titre <sup>a</sup>	ECP	Microscopie épifluorescence	
						VP1	ARNdb
1	CF190025_FRA11	CV-A2	France	7,84	<u>Négatif</u>		Négatif
2	CF179114_FRA11	CV-A2	France	5,86	<u>Négatif</u>		
3	CF192073_FRA11	CV-A2	France	5,94	<u>Négatif</u>	Positif	
4	CF181076_FRA11	CV-A2	France	6,58	<u>Négatif</u>		
5	CF146007_FRA11	CV-A2	France	5,73	<u>Négatif</u>		
6	STU80_DEU05	CV-A2	Allemagne	6,04	<u>Négatif</u>		Négatif
7	CF092105_FRA14	CV-A3	France	ND	<u>Négatif</u>		
8	CF174020_FRA11	CV-A4	France	5,06	<u>Positif</u>		Positif
9	CF188078_FRA12	CV-A4	France	6,32	<u>Négatif</u>		
10	CF063006_FRA11	CV-A4	France	6,01	<u>Négatif</u>		
11	CF177027_FRA12	CV-A4	France	6,36	<u>Négatif</u>		
12	STU24_DEU10	CV-A5	Allemagne	5,47	<u>Négatif</u>	Positif	
13	CF193056_FRA11	CV-A5	France	6,04	<u>Négatif</u>		
14	CF147012_FRA10	CV-A6	France	6,13	<u>Positif</u>		Positif
15	CF175075_FRA10	CV-A6	France	6,27	<u>Négatif</u>		Positif
16	CF325070_FRA12	CV-A6	France	4,48	<u>Négatif</u>		
17	So2640_FIN08	CV-A6	Finlande	ND	<u>Négatif</u>		
18	STU22_DEU09	CV-A6	Allemagne	5,34	<u>Négatif</u>		
19	CF176008_FRA11	CV-A8	France	6,25	<u>Positif</u>		Positif
20	CF179052_FRA11	CV-A8	France	5,8	<u>Positif</u>		
21	CF218013_FRA10	CV-A8	France	5,1	<u>Négatif</u>		
22	CF173027_FRA11	CV-A8	France	ND	<u>Négatif</u>		
23	CF203060_FRA10	CV-A10	France	6,54	<u>Négatif</u>		
24	CF204025_FRA10	CV-A10	France	7,2	<u>Positif</u>		Positif
25	CF276015_FRA12	CV-A10	France	6,73	<u>Négatif</u>		
26	CF209018_FRA12	CV-A10	France	6,88	<u>Négatif</u>		
27	RP96_AUT04	CV-A14	Autriche	ND	<u>Positif</u>	Positif	Positif
28	1062_ISL03	CV-A14	Islande	ND	<u>Positif</u>		Positif
29	E2202_FIN07	CV-A14	Finlande	ND	<u>Négatif</u>		
30	CF279014_FRA10	CV-A16	France	3,47	<u>Positif</u>		Positif
31	BUD25_HUN08	CV-A16	Hongrie	ND	<u>Positif</u>	Positif	Positif
32	BER53_2_GER03	CV-A16	Allemagne	ND	<u>Positif</u>	Positif	Positif
33	CF350028_FRA11	CV-A16	France	3,59	<u>Positif</u>		Positif
34	CF210042_FRA06	EV-A71/C1	France	3,9	<u>Positif</u>		Positif
35	STU546711_DEU08	EV-A71/C1	Allemagne	ND	<u>Positif</u>		Positif
36	STU73_DEU07	EV-A71/C2	Allemagne	5	<u>Négatif</u>	Positif	Positif
37	CF280001_FRA11	EV-A71/C2	France	ND	<u>Positif</u>		Positif
38	CF160019_FRA10	EV-A71/C2	France	5,11	<u>Positif</u>	Positif	Positif
39	STU85_DEU05	EV-A71/C2	Allemagne	ND	<u>Positif</u>	Positif	Positif
40	CF1113_FRA00	EV-A71/C2	France	5,14	<u>Positif</u>	Positif	Positif
41	Bud38HUN04	EV-A71/C4	Hongrie	4,36	<u>Positif</u>	Positif	Positif
42	CF282_FRA97	E-30	France	6,5	<u>Positif</u>		Positif

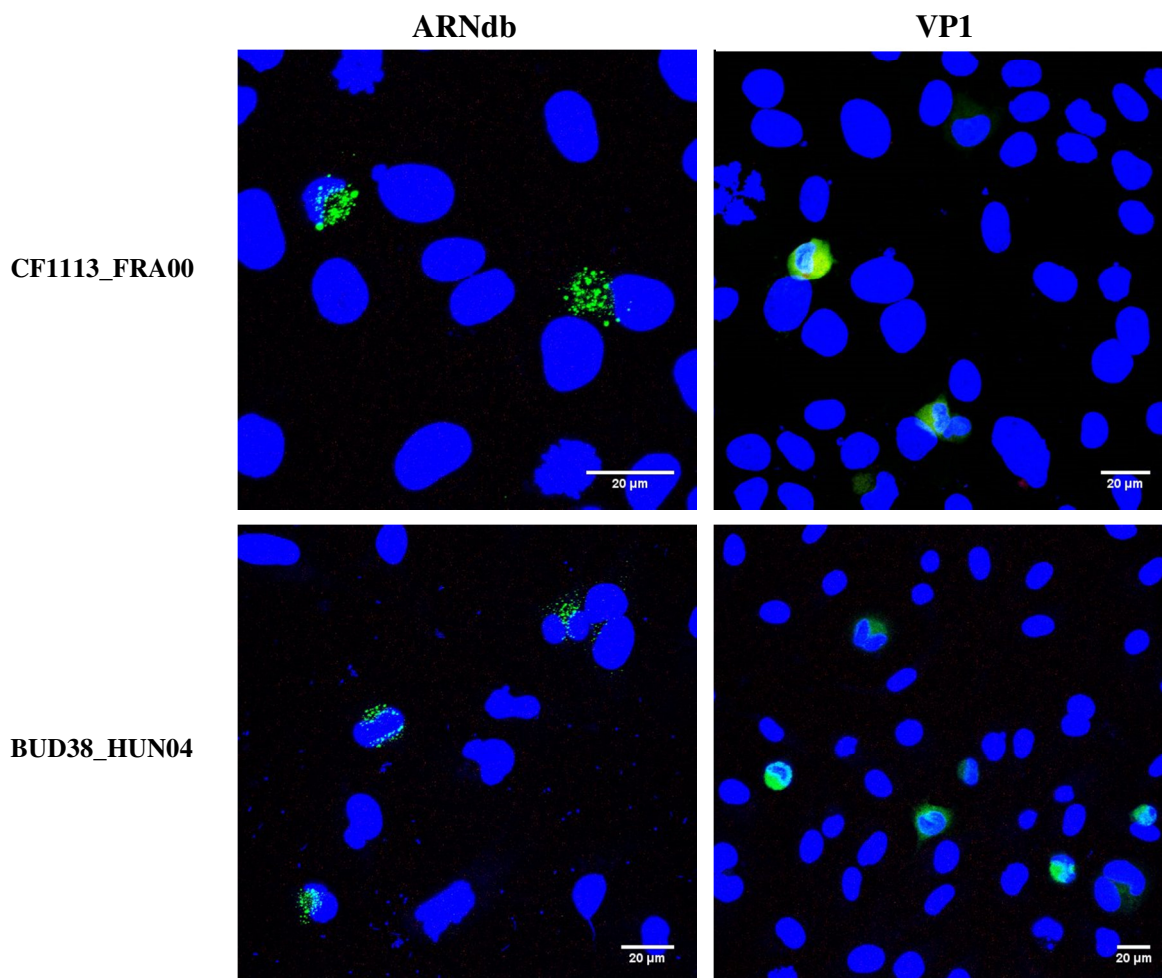
<sup>a</sup> Titre infectieux du stock viral utilisé, exprimé en nombre log du nombre de particules pour 50 µL

## 2) Sensibilité de la lignée HMEC-1 à l'infection par différents EV-A

Un échantillon de 41 souches d'EV-A choisies parmi tous les sérotypes disponibles a été utilisé pour tester la sensibilité de la lignée HMEC-1 à l'infection : CV-A2 (n=6), CV-A3 (n=1), CV-A4 (n=4), CV-A5 (n=2), CV-A6 (n=5), CV-A8 (n=4), CV-A10 (n=4), CV-A14 (n=3), CV-A16 (n=4) et EV-A71 (n=8). Ces tests ont consisté à déterminer l'intensité de la lyse d'une monocouche cellulaire et de la multiplication virale par la technique de microscopie à fluorescence. A 48 heures post inoculation, les ECP ont été observés en microscopie optique. Les souches qui présentent des propriétés lytiques appartiennent principalement à quatre sérotypes : CV-A8, CV-A14, CV-A16 et EV-A71 [Tableaux 18]. La souche témoin d'E-30 (EV-B) appartient aussi à ce groupe. Les autres souches appartenant aux six autres sérotypes induisent peu ou aucune lyse cellulaire à l'exception des trois souches suivantes : CF174020\_FRA11 (CV-A4), CF174012\_FRA10 (CV-A6) et CF204025\_FRA10 (CV-A10).

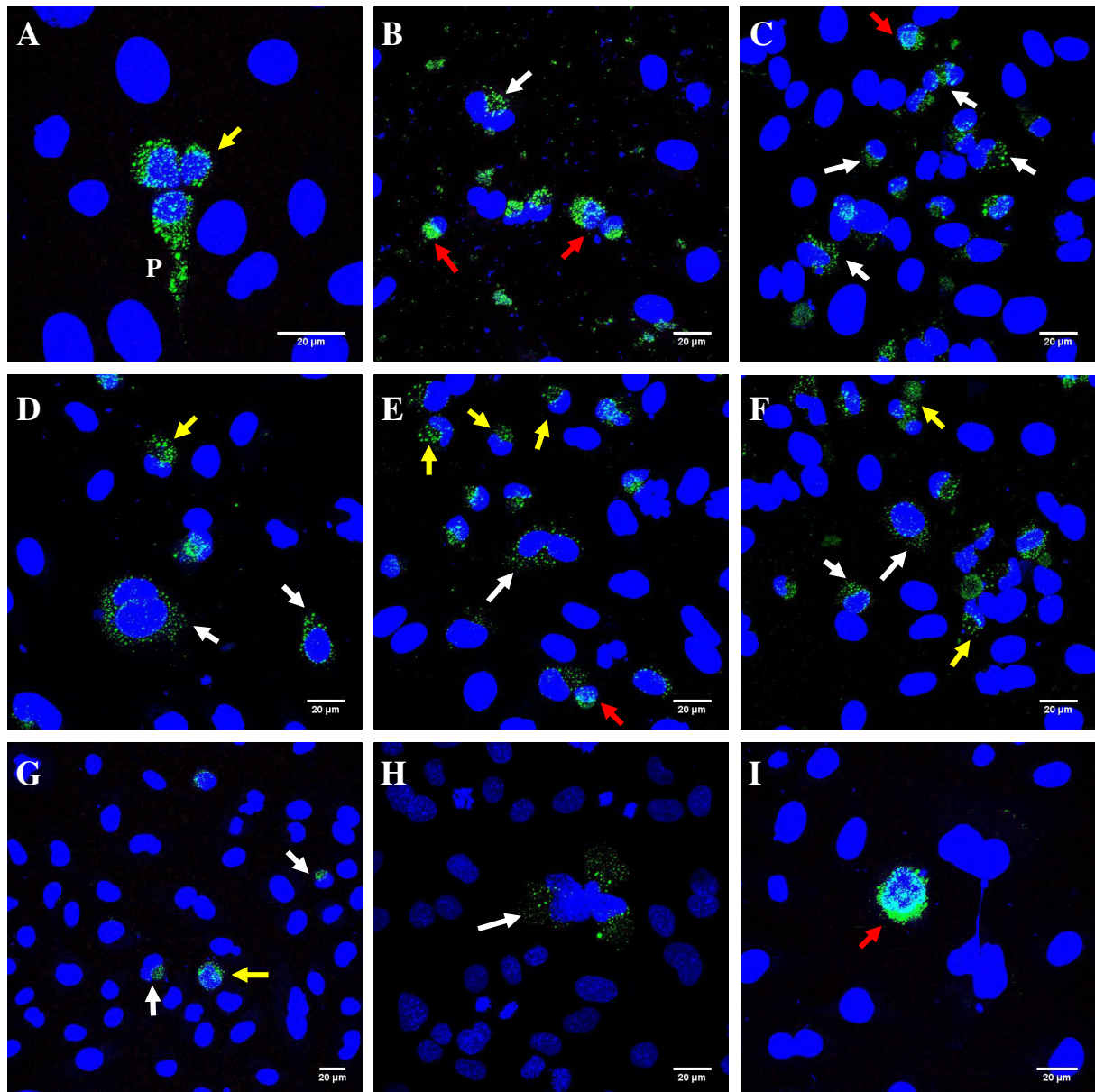
Les ECP ont été vérifiés par microscopie à épi fluorescence à l'aide de marqueurs permettant de caractériser l'infection virale par la présence d'ARN double brin viral et de la protéine de capsid VP1 [Tableau 18]. Les résultats confirment les observations d'ECP pour l'ensemble des 42 souches étudiées à l'exception de quatre d'entre elles. Bien qu'aucun ECP n'ait été observé pour les souches CF192073\_FRA11 (CV-A2), STU24\_DEU10 (CV-A5), CF175075\_FRA10 (CV-A6) et STU73\_DEU07 (EV-A71), la détection de la protéine de capsid VP1 indique une infection virale. La qualité du marquage fluorescent obtenu avec les deux marqueurs a été comparée par microscopie confocale. Le marquage de la protéine de capsid VP1 se caractérise par une fluorescence diffuse au sein du cytoplasme des cellules infectées [Figure 55]. Au contraire, le marquage de l'ARNdb se présente sous la forme de foyers individualisés et délimités très finement.

L'analyse en microscopie confocale a permis de caractériser qualitativement trois stades de l'infection virale à l'échelle intracellulaire pour neuf souches lytiques (CV-A14, n=1 ; CV-A16, n=2 et EV-A71, n=6) : (i) un stade précoce caractérisé par des spots fluorescents de petite taille répartis dans le cytoplasme ; (ii) un stade intermédiaire représentant le début de la multiplication virale se manifestant par une intensité de fluorescence et (iii) un stade de fin d'infection accompagné d'une rétraction cellulaire et d'une réplication péri nucléaire. Ces trois stades sont observés pour la souche STU546711\_DEU08 (EV-A71/C1) [Figure 56E]. Le stade précoce (flèches blanches) est souvent retrouvé sur les différentes observations notamment pour les souches



**Figure 55. Comparaison de deux marqueurs de l'infection virale, la protéine de capsid VP1 et l'ARNdb en microscopie confocale pour deux souches d'EV-A71**

Les cellules HMEC-1 ont été infectées par deux souches d'EV-A71, CF1113\_FRA00 et BUD38\_HUN04 et observées au bout de 48 heures de culture. La fluorescence verte caractérise le marquage soit de la protéine de capsid VP1 soit de l'ARN double brin (ARNdb). Les noyaux des cellules apparaissent en bleu (marquage avec la molécule Hoechst 33342).



**Figure 50. Caractéristiques de l'infection des cellules de la lignée HMEC-1 par des souches de différents sérotypes d'EV-A**

Les cellules HMEC-1 ont été inoculées avec des souches de différents sérotypes de l'espèce EV-A et observées en microscopie confocale après 48 heures de culture : **A**, CV-A14 (RP96\_AUT04) ; **B**, CV-A16 (CF279014\_FRA10) ; **C**, CV-A16 (CF350028\_FRA11) ; **D**, EV-A71/C1 (CF210042\_FRA06) ; **E**, EV-A71/C1 (STU546711\_DEU08) ; **F**, EV-A71/C2 (CF280001\_FRA11) ; **G**, EV-A71/C2 (CF160019\_FRA10) ; **H**, EV-A71/C2 (CF1113\_FRA00) et **I**, EV-A71/C4 (BUD38\_HUN04). La fluorescence verte correspond à l'ARN double brin (ARNdb), un marqueur de l'infection virale. Les noyaux sont marqués en bleu avec la molécule Hoechst 33342. Les flèches de différentes couleurs indiquent des stades qualitatifs de l'infection virale : stade précoce (blanc), stade intermédiaire (jaune) et stade de fin d'infection (rouge). **P**, prolongement cytoplasmique.



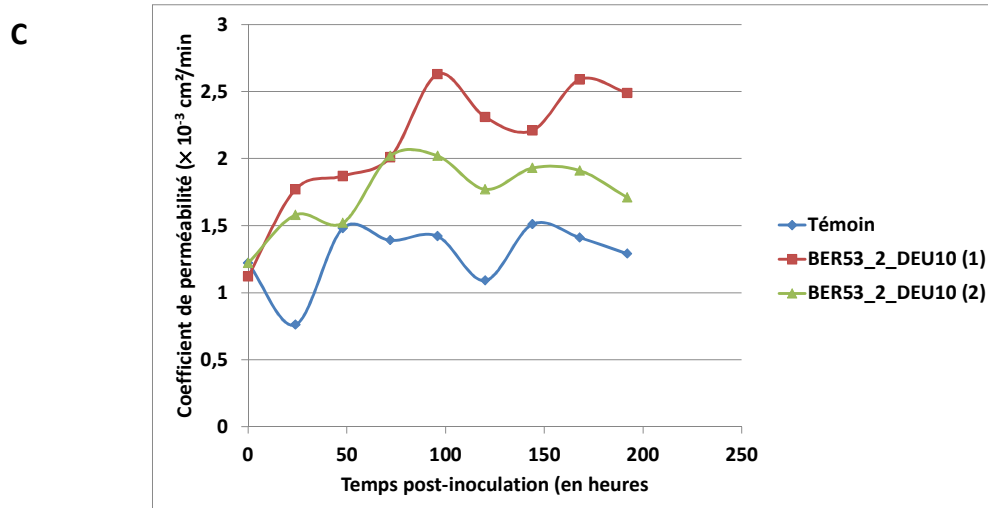
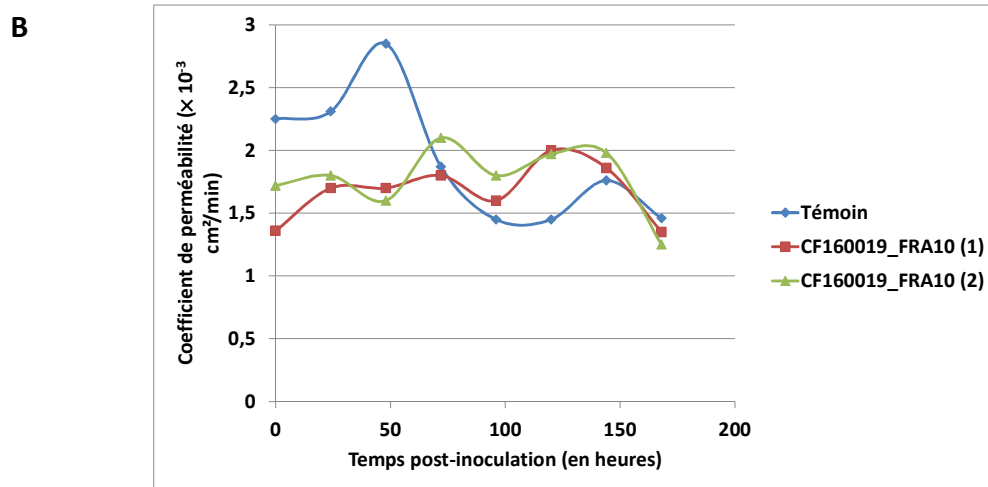
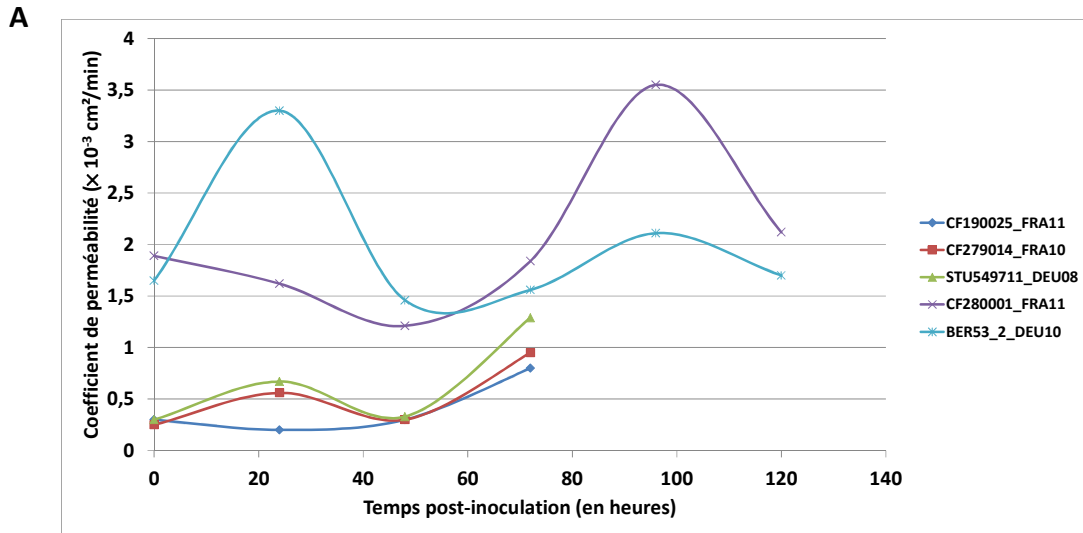
CF350028\_FRA11 (CV-A16), CF210042\_FRA06 (EV-A71/C1), CF280001\_FRA11 (EV-A71/C2) et CF1113\_FRA00 (EV-A71/C2) [Figures 56C, 56D, 56F, 56H]. Le stade intermédiaire se distingue par une condensation du noyau de la cellule hôte, visible particulièrement pour la souche STU546711\_DEU08 (EV-A71/C1) [Figure 56E]. Le stade de fin d'infection se retrouve sur l'observation microscope réalisée pour la souche BUD38\_HUN04 (EV-A71/C4), caractérisée par une rétractation cellulaire et une importante réplication virale péri nucléaire [Figure 56I]. L'observation de la souche RP96\_AUT04 (CV-A14) indique la présence d'un prolongement cytoplasmique riche en ARN viral de la cellule hôte en cours de condensation [Figure 56A].

### 3) Développement d'un modèle de barrière endothéliale de derme

Les cellules HMEC-1 ont été inoculées après avoir été cultivées à la surface d'une membrane microporeuse afin d'obtenir une monocouche étanche et l'étanchéité paracellulaire a été mesurée par un test de clairance du fluorophore «Lucifer Yellow». L'étanchéité paracellulaire est atteinte lorsque le coefficient de perméabilité est inférieur à  $2 \times 10^{-3}$  cm<sup>2</sup>/min.

Une série de tests préliminaires a été réalisée avec 5 souches d'EV-A (1 CV-A2, 2 CV-A16 et 2 EV-A71). Pour les souches CF190025\_FRA11 (CV-A2), CF279014\_FRA10 (CV-A16) et STU549711\_DEU08 (EV-A71/C1), la perméabilité paracellulaire ne dépasse pas  $1,29 \times 10^{-3}$  cm<sup>2</sup>/min durant toute la mesure de cinétique entre 0 et 72 h post infection [Figure 57A]. Pour la souche de CV-A16, BER53\_2\_DEU10, le coefficient de perméabilité atteint un sommet à  $3,3 \times 10^{-3}$  cm<sup>2</sup>/min à 24h post infection puis diminue à 1,46. Par la suite, il est proche de  $2 \times 10^{-3}$  cm<sup>2</sup>/min jusqu'à 120h post infection. L'infection de l'insert par la souche d'EV-A71, CF280001\_FRA11, induit une perméabilité ( $3,55 \times 10^{-3}$  cm<sup>2</sup>/min) à 96 h post infection.

Afin de vérifier ces tendances de perméabilité, deux cinétiques d'infection ont été réalisés en utilisant une barrière témoin et deux barrières infectées par une souche virale d'EV-A. La première cinétique a été faite avec une souche d'EV-A71/C2, CF160019\_FRA10, sur 168h. L'insert témoin présente une forte augmentation de la valeur de perméabilité à 48h suivi d'une diminution progressive jusqu'à environ  $1,5 \times 10^{-3}$  cm<sup>2</sup>/min [Figure 57B]. La perméabilité des deux autres inserts infectés par la souche virale varie entre  $1,5 \times 10^{-3}$  et  $2,1 \times 10^{-3}$  cm<sup>2</sup>/min durant toute cette cinétique. La seconde cinétique a été réalisée avec une souche très lytique pour la lignée HMEC-1, BER53\_2\_DEU03 (CV-A16) durant 192h [Figure 57C].



**Figure 52.** Cinétique de la perméabilité cellulaire des cellules HMEC-1 infectées par les souches d'EV-A indiquées

La perméabilité de la barrière témoin reste globalement stable à environ  $1,4 \times 10^{-3}$  cm<sup>2</sup>/min durant toute la cinétique. La perméabilité des barrières infectées par la souche virale augmente au cours du temps pour atteindre respectivement les valeurs  $2 \times 10^{-3}$  et  $2,5 \times 10^{-3}$  cm<sup>2</sup>/min, indiquant une certaine perméabilité du tapis cellulaire pour le fluorophore. Cette souche de CV-A16 entraîne une destruction de l'intégrité de la barrière endothéliale microvasculaire de derme.

# **DISCUSSION**

La généralisation du génotypage pour identifier les entérovirus a considérablement accru les données génomiques des souches associées aux manifestations cliniques et il est devenu envisageable de relier les données de parenté génétiques aux autres informations disponibles : caractéristiques cliniques et biologiques des infections, informations sur l'origine géographique et temporelle des souches. L'analyse croisée des données a permis des avancées scientifiques décisives dans la compréhension d'infections causées par les virus VIH, influenza et VHC. Les méthodes qui permettent de traiter ces données se rattachent à trois types d'approches : (1) l'épidémiologie basée sur des modèles à compartiments, (2) la génétique des populations et (3) l'analyse phylogénétique. Pour analyser de façon globale l'épidémiologie des sérotypes de l'espèce EV-A, nous avons choisi la troisième approche en utilisant les méthodes phylogénétiques les plus récentes.

## **I. L'épidémiologie virale analysée à partir de données moléculaires**

Le logiciel BEAST a été conçu pour étudier l'épidémiologie moléculaire des virus à ARN par une approche bayésienne. L'histoire évolutive et la dissémination géographique ont ainsi pu être reconstruites pour plusieurs virus émergents : influenza H5N1 (Lam *et al*, 2012), influenza H1N1 (Rambaut *et al*, 2008 ; Smith *et al*, 2009) ou le virus de la dengue (Allicock *et al*, 2012). Une étude récente, reprise dans de nombreux articles de presse, a établi l'origine de la pandémie mondiale de VIH à partir d'un échantillon de plus de 800 séquences anciennes (Faria *et al*, 2014).

Le logiciel BEAST a été relativement peu utilisé pour étudier l'épidémiologie moléculaire des entérovirus. Les premières études ont permis de dater l'émergence de variants parmi des sérotypes de l'espèce EV-B (McWilliam Leitch *et al*, 2009 ; McWilliam Leitch *et al*, 2010 ; Gullberg *et al*, 2010 ; Bailly *et al*, 2011). Notre équipe a utilisé le logiciel BEAST afin d'approfondir le lien entre les processus évolutifs du sérotype CV-B5 et la transmission de ce virus en Europe (Henquell *et al*, 2013).

## **II. Etudes antérieures de l'histoire évolutive de l'EV-A71**

La première étude Bayésienne concernant l'EV-A71 a utilisé un échantillon de 628 séquences (génogroupe A, n=1 ; génogroupe B, n=283 ; génogroupe C, n=344) (Tee *et al*, 2010). Elle avait suggéré des transports de souches entre des zones géographiques distantes sans en

apporter de preuve. L'émergence du virus avait été datée en 1941 et l'histoire démographique avait suggéré un lien entre les pics de diversité génétique et les épidémies.

Concernant l'Europe, une première étude avait concerné les souches d'EV-A71 en circulation aux Pays-Bas entre 1963 et 2008 (Van der Sanden *et al*, 2010). Elle avait suggéré que les génogroupes B et C de l'EV-A71 avaient des histoires évolutives différentes : un remplacement rapide des souches au cours du temps pour le génogroupe B et une co-circulation de sous-populations (C1 et C2) pour le génogroupe C. Le profil démographique indique un pic de forte incidence en 1986 correspondant à une épidémie causée par des souches B2 (Van der Sanden *et al*, 2010). La seconde étude, réalisée par notre équipe, avait analysé un échantillon de 848 séquences du locus 1D<sup>VP1</sup> (dont 86 souches européennes ; Mirand *et al*, 2010). Les profils phylogénétiques avaient confirmé les différences d'histoire évolutive des génogroupes B et C à une échelle plus globale. Ces travaux suggéraient qu'en Europe, les sous génogroupes C1 et C2 co-circulaient sous la forme de deux populations virales. Les chronogrammes indiquaient un lien phylogénétique entre les populations virales européennes et asiatiques et suggéraient une dissémination sur de grandes distances entre l'Europe et l'Asie (Mirand *et al*, 2010).

### **III. Apport de la dynamique géographique à l'épidémiologie de l'EV-A71**

Notre étude, réalisée avec un nombre important de séquences génomiques ( $n > 1675$  ; locus 1D<sup>VP1</sup>), est la première à avoir employé des modèles phylogéographiques explicites pour analyser la diffusion des souches et leur circulation à une échelle géographique globale (intercontinentale Europe/Asie) ou entre des pays proches (échelle régionale).

La topologie en escalier des arbres phylogénétiques des populations d'EV-A71 indique un processus d'extinction et de remplacement rapide des lignées au cours du temps [Figure 40], ce qui confirme les hypothèses antérieures (Tee *et al*, 2010). Les profils démographiques estimés à partir des données moléculaires suggèrent que les pics épidémiques sont rares à une échelle globale [Figure 41]. L'histoire évolutive de chaque sous-génogroupe d'EV-A71 présente des spécificités en termes de fréquence des épidémies et de distribution géographique. Les virus B5 et C4 se caractérisent par une transmission épidémique en Asie du sud alors qu'en Europe, les virus variants C1 et C2 ont une transmission discrète, la seule épidémie récente s'est produite en 2007 aux Pays-Bas (Van der Sanden *et al*, 2009).

L'analyse phylogéographique des données moléculaires du variant C2 n'est pas en faveur d'une source virale unique pour cette épidémie : elle suggère au contraire de multiples introductions virales qui ont aussi affecté d'autres pays Européens (Allemagne, France) sans pour autant y produire d'épidémie. Ces résultats conduisent à l'hypothèse que des facteurs locaux (incluant des facteurs propres aux populations d'individus) déterminent le « succès épidémique » des introductions sporadiques des souches d'entérovirus. Ce profil épidémiologique caractérisé par des vagues de transmissions sporadiques en Europe a été retrouvé ultérieurement en 2010 et en 2013. L'origine précise de ces introductions sporadiques est difficile à retracer : un biais de l'échantillonnage rend plus difficile l'estimation des transmissions de souches entre régions distantes (Europe / Asie). L'étude phylogéographique permet de déterminer les événements de transmission principalement entre les pays européens proches et montre l'importance des flux de virus sur de courtes distances qui avaient été suggérés dans des travaux antérieurs (Mizuta *et al*, 2005 ; Mirand *et al*, 2010).

Le virus EV-71/C4 circule très massivement en Asie et plusieurs cas d'infections causées par ce variant ont été rapportés pour la première fois dans plusieurs pays européens en 2004. Notre étude a montré que les souches ont constitué une chaîne de transmission entre trois régions géographiques : Asie (Japon), Europe (Allemagne, Croatie, France, Hongrie) et Amérique du nord (Canada) pendant une période de 5 ans (2002–2006). La transmission des souches C4 s'est produite en Europe alors que les cas rapportés d'infections à virus C1 et C2 étaient très peu fréquents. Récemment, d'autres infections sporadiques à virus C4 ont été rapportées hors du bassin asiatique de ce variant, en France (Schuffenecker *et al*, 2014), au Danemark (Fischer *et al*, 2014) et en Russie [Figure 40]. Les séquences génomiques de ces souches n'ont pas de relation phylogénétique directe avec celles de la chaîne de transmission, ce qui suggère des introductions virales récentes dont l'origine géographique exacte ne peut pas être retracée. A partir de ces données, on peut formuler l'hypothèse que l'EV-A71/C4 peut être introduit en Europe lorsque la transmission des autres virus variants demeure très faible pendant plusieurs années consécutives. Cette alternance des virus variants d'EV-A71 en Europe pourrait résulter de l'immunité croisée induite par les souches génétiquement proches. En effet, l'analyse comparée de leurs variations antigéniques par séroneutralisation est en faveur d'épitopes communs entre les souches issues des différents sous-génogroupes C (Huang *et al*, 2013).

#### IV. Dynamique des populations virales et épidémiologie des autres EV-A

Les épidémies de maladie pied-main-bouche et d'herpangine causées par les sérotypes CV-A6 et CV-A10 dans plusieurs pays européens ont mis en lumière deux sérotypes d'EV-A qui étaient jusqu'à présent considérés comme moins fréquents que le CV-A16 et l'EV-A71 (Blomqvist *et al*, 2010). La surveillance des cas maladie PMB et d'herpangine initiée par le laboratoire en 2010 a permis de montrer la circulation épidémique des types CV-A6 et CV-A10 dans la population pédiatrique de la région de Clermont-Ferrand (Mirand *et al*, 2012). La banque de souches constituée nous a permis d'accroître considérablement le nombre des séquences génomiques pour ces deux entérovirus. Il existe cependant un biais dans l'échantillon de séquences car deux pays (la Chine et la France) sont surreprésentés. Notre étude d'épidémiologie phylogénétique des coxsackievirus CV-A6, CV-A10 et CV-A16 est la première à faire appel aux méthodes bayésiennes. Une étude phylogéographique basée sur un modèle de caractères discrets a été réalisée avec un échantillon de 58 séquences de CV-A4 (partie 3' du locus 1D<sup>VP1</sup> ; Chu *et al*, 2011).

La structure phylogénétique du CV-A6 indique une diversité inhabituelle puisque le chronogramme montre qu'il existe huit clades ou sous-génogroupes. L'épidémie identifiée dans la région de Clermont-Ferrand en 2010 a été causée par un variant (Figure S3) dont l'ancêtre commun a émergé en 2002 [1999–2004]. La phylogénie du CV-A6 montre aussi une co-circulation à l'échelle géographique d'un pays (France ou Chine) de souches appartenant à de multiples sous-génogroupes. On retrouve des caractéristiques phylogénétiques semblables pour le CV-A10. L'épidémiologie de ce virus est caractérisée par la co-circulation à une échelle géographique globale de 5 populations virales. Le sixième génogroupe est représenté uniquement par la souche prototype. Les souches de CV-A10 isolées lors de l'épidémie de 2010 en France se répartissent en différents groupes monophylétiques. Comme pour le CV-A6, cette caractéristique suggère que l'épidémie a eu des origines (importations) multiples ; il n'existe pas de souche d'origine asiatique dans le groupe de l'épidémie, mais il peut s'agir d'un biais d'échantillonnage.

L'étude d'un grand échantillon de séquences de CV-A16 par une approche phylogénétique « Neighbor Joining » a montré l'existence de deux génogroupes (Chen *et al*, 2013). Notre étude a consolidé l'échantillonnage avec 157 nouvelles séquences du locus 1D<sup>VP1</sup> déterminées chez des souches isolées en Europe (France, Allemagne et Hongrie ; 2002–2014). La structure phylogénétique estimée avec cet échantillon montre que la très



grande majorité des infections à CV-A16 rapportées en Europe et en Chine a été causée par des souches du sous-génogroupe B3. Le résultat principal de notre étude est la mise en évidence d'un nouveau génogroupe que nous avons désigné par la lettre C. Le génogroupe C est représenté par une trentaine de souches isolées en France entre 2011 et 2014, ce qui indique la persistance de sa circulation. Il a émergé en 2004 [2001 – 2007] soit 7 ans avant sa diffusion en France. Une séquence issue de la base de données GenBank (KF956720) est phylogénétiquement apparentée à ce nouveau génogroupe : la souche correspondante a été isolée au Pérou en 2009. Sur la branche phylogénétique du génogroupe C, cette « séquence péruvienne » est localisée en amont des « séquences françaises », ce qui suggère une diffusion géographique qu'il n'est pas possible de retracer à cause du biais d'échantillonnage.

## **V. Apport de la phylodynamique à l'épidémiologie des maladies virales**

Il n'existe pas de réseau sentinelle en Europe pour la surveillance de la maladie PMB et la circulation de l'EV-A71. Cette maladie est présente en Europe, comme le montre les données du réseau de surveillance mise en place à l'échelle de la France par notre laboratoire. Les épidémies de maladie PMB sont causées majoritairement par des EV autres que l'EV-A71 dont la circulation demeure sporadique à l'échelle européenne.

Avec la découverte des génogroupes E, F et G en Afrique et en Inde, plusieurs études récentes ont montré que la diversité de l'EV-A71 ne se limitait pas aux génogroupes B et C. L'existence de ces virus soulève plusieurs questions dont celles de leurs caractéristiques neurotropes et de leur circulation réelle sous la forme de populations virales établies. Avec la recrudescence et le développement des transports humains sur de longues distances, il est possible que ces virus variants évoluent et s'adaptent à des niches épidémiologiques.

Un réseau de surveillance des EV-A71 pourrait être mis en place en collaboration avec des laboratoires africains et indiens. Cette initiative permettrait d'accroître considérablement l'échantillonnage des souches de sorte que les méthodes de modélisation à partir des données moléculaires seraient plus fiables. Avec un échantillonnage représentatif, il serait par exemple possible de retracer les flux de la circulation transcontinentale des virus variants.

## **VI. Evolution adaptative de la capsid de l'EV-A71 au cours de la circulation des souches**

Les virus sont en permanence soumis à des pressions sélectives qui réduisent la survie et/ou la reproduction des génotypes. La diffusion des entérovirus, à l'exception des poliovirus, s'effectue dans la population humaine en l'absence de lutte médicamenteuse ou vaccinale et de changement d'espèce hôte. L'évolution adaptative de ces virus est donc soumise à une sélection exercée essentiellement par le système immunitaire de l'hôte et les changements d'environnement cellulaire lors de la multiplication intra-individuelle.

Les changements protéiques estimés dans la région P1 sont différents entre les génogroupes B et C. Les changements adaptatifs, plus nombreux pour le génogroupe B, sont localisés en majorité dans la protéine VP1. Les variations qui affectent les boucles protéiques (reliant les feuillets  $\beta$ ) sont réparties différemment entre les deux génogroupes d'EV-A71 [Figure 46]. Ces sites sont potentiellement très exposés à la surface de la capsid, ce qui suggère un rôle en tant qu'épitopes antigéniques. Cette hypothèse pourrait être examinée par une étude structurale complétée d'une analyse fonctionnelle à l'aide d'anticorps neutralisants.

L'évolution moléculaire du génogroupe B est caractérisée par un plus grand nombre de mutations réverses que celle du génogroupe C, exception faite du variant C4. Les mutations parallèles, mutations qui surviennent aux mêmes sites indépendamment dans des lignées différentes. Ces caractéristiques évolutives reflètent les contraintes structurales importantes qui pèsent sur l'évolution moléculaire de la capsid : le taux de mutations non synonymes sur les mutations synonymes ( $dN/dS$ )  $< 0,041$  pour l'ensemble de la région P1 (Tableau 12). Le taux  $dN/dS$  des populations B et C4 est supérieur à celui des populations C1 et de C2, ce qui suggère la présence de sites soumis à une sélection positive. Parmi ces sites figurent les positions 98 et 145 de la protéine VP1, lesquelles sont impliquées dans la liaison de la capsid au récepteur cellulaire PSGL-1 [Nishimura et al, 2013]. Pour le variant C4, s'ajoutent les positions 92 de VP3 et 249 de VP1 dont le rôle fonctionnel reste à élucider.

Ces différences d'histoire évolutive et de contraintes sélectives pourraient être une conséquence des réponses immunes de l'hôte. Lors de la co-circulation de souches de différents génogroupes, les épitopes qui leur sont communs provoquent des réactions d'immunité croisée, ce qui peut conduire à des changements de populations virales en faveur de celles qui présentent des changements adaptatifs d'échappement. Lorsque la population hôte, sensible au virus est très réduite par l'installation d'une mémoire immunitaire, les virus disposant d'épitopes différents pourront être sélectionnés positivement et infecter de

nouveaux individus. La protéine VP1 est une protéine clé de l'histoire évolutive et structurale de l'EV-A71. Son exposition à la surface de la capsid est à l'origine d'une intense réponse immunitaire humorale qui en s'exerçant à l'échelle de la population humaine peut conduire à une évolution adaptative. *A contrario*, le rôle de la protéine dans le tropisme cellulaire limite considérablement les possibilités de cette évolution adaptative des sites antigéniques.

La localisation structurale des mutations mises en évidence par les analyses phylogénétiques de l'EV-A71 est en cours de validation en utilisant les structures cristallographiques disponibles pour ce virus (Wang *et al*, 2012). Cette étude permettra de mieux comprendre leur implication dans l'évolution virale et la persistance des populations au cours du temps. Les cartes de localisation structurale ainsi établies ouvriront la voie à des études fonctionnelles. Une approche consisterait à générer des virus chimères comportant des mutations dans les sites répertoriés et à évaluer le pouvoir infectieux avec un modèle animal ou un modèle cellulaire approprié (cf. dernière partie de la discussion). Les sites exposés pourraient aussi être comparés par neutralisation croisée afin d'analyser le lien entre le pouvoir infectieux et la structure antigénique.

## VII. Recombinaison génétique et évolution des EV-A

Chez les picornavirus, la recombinaison génétique est reliée à l'émergence de variants viraux dont certains ont été associés à des symptômes cliniques sévères. Par exemple, il a été montré que des souches vaccinales de poliovirus issues d'une recombinaison avec des souches de coxsackievirus A de l'espèce EV-C étaient neuropathogènes (Combelas *et al*, 2011). Pour autant, le lien de causalité entre recombinaison et acquisition de propriétés neurovirulentes n'est pas prouvé. La recombinaison intertypique est fréquente parmi les EV-B mais n'est pas reliée directement à la virulence. Le rôle de la recombinaison chez les entérovirus est d'autant plus difficile à cerner que sa fréquence varie parmi les espèces.

Nous avons analysé un échantillon de génomes d'EV-A car il a été montré que la recombinaison intertypique est peu fréquente dans ce groupe (Simmonds et Welch, 2006). Cette analyse de 394 souches a une valeur indicative, car elle a permis d'identifier 22 souches ayant une origine recombinante intertypique probable, soit 5,6% de l'échantillon. Quatre événements de recombinaison intratypique ont aussi été identifiés parmi les séquences de notre échantillon (CV-A4, CV-A6, CV-A10 et CV-A16). Il n'y a pas de différence majeure entre les sérotypes retrouvés dans les deux types de recombinaison. Ces données confirment la rareté des échanges génomiques parmi les EV-A.

Les évènements de recombinaison identifiés chez le type CV-A16 concernent le génogroupe C dont l'émergence semble liée à un évènement de recombinaison intertypique. La souche génétiquement la plus proche dans notre échantillon est un CV-A10 (souche isolée en Russie) mais il n'est pas possible d'affirmer un échange génétique direct car la similitude n'est que de 90%. Nos données confirment l'hypothèse selon laquelle l'émergence d'un génogroupe est liée à une recombinaison avec un autre sérotype du pool global des EV-A (Lukashev *et al*, 2014). L'identification d'un évènement de recombinaison intratypique parmi les souches du génogroupe C est particulièrement intéressante parce qu'elle s'est produite avec le génogroupe B3. La similitude de 98% avec une souche isolée en France suggère un évènement récent et une circulation active du nouveau virus variant. D'après les données épidémiologiques du CNR Entérovirus/Parechovirus, les infections à génogroupe C sont liées aux mêmes manifestations cliniques que celles causées par les autres CV-A16.

Le type EV-A71 présente une circulation intense en Asie où il est en co-circulation avec de nombreux autres types (Solomon *et al*, 2010). Malgré ces facteurs favorables à la recombinaison, nous n'en avons pas trouvé de preuve parmi les souches de notre échantillon. L'absence de recombinaison intertypique suggère que ce mécanisme génétique a un rôle mineur parmi les populations virales actives sur le plan épidémiologique (dissémination des souches d'EV-A). L'hypothèse selon laquelle la recombinaison est reliée à l'émergence des certains sous-génogroupes d'EV-A71 est toutefois vraisemblable (McWilliam Leitch *et al*, 2010). Des sites de recombinaison ont été localisés dans le gène 2A chez les souches des sous-génogroupes B3 et C4 confirmant des études précédentes (Chan et Abubakar, 2004 ; Van der Sanden *et al*, 2011 ; McWilliam Leitch *et al*, 2012). Les séquences parentales précises n'ont pas été identifiées formellement.

Une limite importante de notre étude de la recombinaison tient au fait que nous avons principalement analysé des portions subgénomiques de sorte que nous ne pouvons pas exclure que des évènements de recombinaison soient présents ailleurs dans les génomes. Il serait nécessaire d'envisager de mettre en place les outils pour effectuer une analyse systématique des génomes complets à partir des échantillons ou des souches isolées en culture. Afin d'approfondir le rôle de la recombinaison génétique dans l'évolution des EV-A, une approche fonctionnelle par réalisation de virus chimères est envisageable. Des virus recombinants entre différents sérotypes ou génogroupes permettraient d'analyser les caractéristiques phénotypiques (pouvoir infectieux, stabilité du génome). Cette approche pour générer des virus chimères d'EV-A permettrait par exemple d'explorer l'hypothèse selon laquelle l'EV-A71 ou le CV-A16 sont réfractaires à la recombinaison intertypique.

## VIII. Développement d'un modèle d'étude du pouvoir pathogène des EV-A

La physiopathologie des infections cutanées et muqueuses causées par les EV-A est peu explorée car les modèles cellulaires ou animaux sont inadaptés. Nous avons initié la mise au point d'un modèle qui servirait à décrire les processus physiopathologiques de la maladie pied-main-bouche ou de l'herpangine. Ce modèle fait appel à la lignée cellulaire endothéliale microvasculaire de derme « HMEC-1 » (Ades *et al*, 1992). L'utilisation de cette lignée s'est révélée difficile à cause d'une croissance cellulaire rapide et de perte de l'inhibition de contact. D'après les données de la littérature, la sénescence cellulaire n'est pas observée avant le passage 95 (Ades *et al*, 1992).

Nous avons montré que les cellules HMEC-1 sont sensibles à l'infection par différents EV-A en utilisant un échantillon de 41 souches isolées au cours d'un syndrome PMB. Une étude antérieure avait montré le pouvoir infectieux de l'EV-A71 dans les cellules HMEC-1 (Liang *et al*, 2004). Notre étude a établi la permissivité de la lignée pour trois sérotypes principalement (EV-A71, CV-A14 et CV-A16). Or, ces trois sérotypes sont génétiquement apparentés et forment un groupe monophylétique [Figure 16]. Il est possible que cette parenté soit issue d'une caractéristique commune héritée de l'ancêtre commun à ces sérotypes : l'utilisation du récepteur membranaire SCARB2. Cette molécule sert de récepteur cellulaire aux souches de ces trois sérotypes lors de l'infection des cellules RD (Yamayoshi *et al*, 2012). Par ailleurs, la permissivité des cellules HMEC-1 à l'infection par l'EV-A71 ne dépend pas du sous-génogroupe. Il existe cependant des variations notables de pouvoir infectieux entre les souches des autres sérotypes d'EV-A. Une étude suggère que le pouvoir infectieux de la souche prototype G-10 du CV-A16 (chez le souriceau nouveau-né) est différent de celui des souches en circulation (Wei *et al*, 2014). Des variations notables de pouvoir infectieux entre des souches de même sérotypes ont été mises en évidence avec un modèle de barrière hémato-encéphalique (Volle *et al*, 2015).

La vitesse de croissance de la lignée HMEC-1 (confluence atteinte en 3 jours) impose un suivi strict de l'évolution du tapis cellulaire de l'insert avec un test quotidien de clairance de fluorophore. Les différentes cinétiques d'infection avec différentes souches présentent des résultats différents qui nécessitent d'avantage d'expérimentations. Il faudrait compléter ces expériences par une mesure de la quantité de génome virale dans les compartiments luminal et abluminal pour les différents tests réalisés par RT-qPCR. Ces expériences pourraient être complétées par une mesure de la résistance électrique transendothéliale. Les conditions de

culture de la lignée HMEC-1 sur une membrane microporeuse doivent être améliorées, en raison de leur vitesse de croissance.

Il s'agit d'un travail préliminaire. Parmi les perspectives à long terme, il serait intéressant de caractériser le(s) récepteur(s) cellulaire(s) impliqué(s) dans l'infection virale de la lignée HMEC-1 par les EV-A. Par ailleurs, les caractéristiques intracellulaires de l'infection virale des EV-A (remodelage cellulaire) pourraient être approfondies par des approches de microscopie (confocale, électronique). A plus long terme, il serait nécessaire de d'utiliser des cultures primaires de cellules endothéliales microvasculaires de derme. L'utilisation des lignées est discutée en raison de l'immortalisation des cellules qui peut induire de multiples changements génétiques et fonctionnels. Par ailleurs, ce modèle pourrait être complété par une co-culture avec des cellules humaines fibroblastiques de derme. Ceci rapprocherait d'avantage le modèle cellulaire *in vitro* des conditions *in vivo* de l'infection de la peau par les EV-A.

# REFERENCES BIBLIOGRAPHIQUES

- AbuBakar S, Sam IC, Yusof J, Lim MK, Misbah S, MatRahim N, Hooi PS. Enterovirus 71 outbreak, Brunei. *Emerg Infect Dis.* 2009. 15(1):79-82.
- Ades EW, Candal FJ, Swerlick RA, George VG, Summers S, Bosse DC, Lawley TJ. HMEC-1: establishment of an immortalized human microvascular endothelial cell line. *J Invest Dermatol.* 1992. 99(6):683-90.
- Agirre A, Barco A, Carrasco L, Nieva JL. 2002. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *J Biol Chem.* 277(43):40434-41.
- Agrez MV, Shafren DR, Gu X, Cox K, Sheppard D, Barry RD. Integrin alpha v beta 6 enhances coxsackievirus B1 lytic infection of human colon cancer cells. *Virology.* 1997. 239(1):71-7.
- Akhmadishina LV, Ereemeeva TP, Trotsenko OE, Ivanova OE, Mikhailov MI, Lukashev AN. Seroepidemiology and molecular epidemiology of enterovirus 71 in Russia. *PLoS One.* 2014. 9(5):e97404.
- Alexander DA, Dimock K. Sialic acid functions in enterovirus 70 binding and infection. *J Virol.* 2002 Nov;76(22):11265-72.
- Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, Mueller BA, Suchard MA, Foster JE, Rambaut A, Carrington CV. Phylogeography and population dynamics of dengue viruses in the Americas. *Mol Biol Evol.* 2012. 29(6):1533-43.
- Alsop J., Flewett Th., Foster Jr. "Hand-foot-and-mouth disease" in Birmingham in 1959. *Br Med J.* 1960. 2(5214):1708-11.
- Andréoletti L, Lévêque N, Boulagnon C, Brasselet C, Fornes P. Viral causes of human myocarditis. *Arch Cardiovasc Dis.* 2009. 102(6-7):559-68.
- Austermann-Busch S, Becher P. RNA structural elements determine frequency and sites of nonhomologous recombination in an animal plus-strand RNA virus. *J Virol.* 2012. 86(13):7393-402.
- Badran SA, Midgley S, Andersen P, Böttiger B. Clinical and virological features of enterovirus 71 infections in Denmark, 2005 to 2008. *Scand J Infect Dis.* 2011. 43(8):642-8.
- Bailly JL, Béguet A, Chambon M, Henquell C, Peigue-Lafeuille H. Nosocomial transmission of echovirus 30: molecular evidence by phylogenetic analysis of the P1 encoding sequence. *J Clin Microbiol.* 2000. 38(8):2889-92.
- Bailly JL, Chambon M, Peigue-Lafeuille H, Laveran H, De Champs C, Beytout D. Activity of glutaraldehyde at low concentrations (less than 2%) against poliovirus and its relevance to gastrointestinal endoscope disinfection procedures. *Appl Environ Microbiol.* 1991. 57(4):1156-60.
- Bailly J-L, Mirand A, Henquell C, Archimbaud C, Chambon M, Charbonné F, Traoré O, Peigue-Lafeuille H. Génotypage et épidémiologie moléculaire des entérovirus non poliomyélitiques. *Virologie.* 2008. 12 (1) : 53-65.
- Bailly JL, Mirand A, Henquell C, Archimbaud C, Chambon M, Regagnon C, Charbonné F, Peigue-Lafeuille H. Repeated genomic transfers from echovirus 30 to echovirus 6 lineages indicate co-divergence between co-circulating populations of the two human enterovirus serotypes. *Infect Genet Evol.* 2011 Mar;11(2):276-89.
- Basavappa R, Syed R, Flore O, Icenogle JP, Filman DJ, Hogle JM. Role and mechanism of the maturation cleavage of VP0 in poliovirus assembly: structure of the empty capsid assembly intermediate at 2.9 Å resolution. *Protein Sci.* 1994. 3(10):1651-69.
- Belnap DM, McDermott BM Jr, Filman DJ, Cheng N, Trus BL, Zuccola HJ, Racaniello VR, Hogle JM, Steven AC. Three-dimensional structure of poliovirus receptor bound to poliovirus. *Proc Natl Acad Sci U S A.* 2000 Jan 4;97(1):73-8.

- Bergelson JM, Chan M, Solomon KR, St John NF, Lin H, Finberg RW. Decay-accelerating factor (CD55), a glycosylphosphatidylinositol-anchored complement regulatory protein, is a receptor for several echoviruses. *Proc Natl Acad Sci U S A*. 1994 Jun 21;91(13):6245-8.
- Bergelson JM, Modlin JF, Wieland-Alter W, Cunningham JA, Crowell RL, Finberg RW. Clinical coxsackievirus B isolates differ from laboratory strains in their interaction with two cell surface receptors. *J Infect Dis*. 1997 Mar;175(3):697-700.
- Bessaud M, Razafindratsimandresy R, Nougairède A, Joffret ML, Deshpande JM, Dubot-Pérés A, Héraud JM, de Lamballerie X, Delpeyroux F, Bailly JL. Molecular comparison and evolutionary analyses of VP1 nucleotide sequences of new African human enterovirus 71 isolates reveal a wide genetic diversity. *PLoS One*. 2014. 9(3):e90624.
- Bible JM, Iturriza-Gomara M, Megson B, Brown D, Pantelidis P, Earl P, Bendig J, Tong CY. Molecular epidemiology of human enterovirus 71 in the United Kingdom from 1998 to 2006. *J Clin Microbiol*. 2008. 46(10):3192-200.
- Bienz K, Egger D, Wolff DA. Virus replication, cytopathology, and lysosomal enzyme response of mitotic and interphase Hep-2 cells infected with poliovirus. *J Virol*. 1973. 11(4):565-74.
- Biswas T. Enterovirus 71 causes hand, foot and mouth disease outbreak in Cambodia. *Natl Med J India*. 2012. 25(5):316.
- Blomberg J, Lycke E, Ahlfors K, Johnsson T, Wolontis S, von Zeipel G. Letter: New enterovirus type associated with epidemic of aseptic meningitis and-or hand, foot, and mouth disease. *Lancet*. 1974. 13;2(7872):112.
- Blomqvist S, Klemola P, Kaijalainen S, Paananen A, Simonen ML, Vuorinen T, Roivainen M. Co-circulation of coxsackieviruses A6 and A10 in hand, foot and mouth disease outbreak in Finland. *J Clin Virol*. 2010. 48(1):49-54.
- Boni MF, Posada D, Feldman MW (2007) An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 176(2): 1035–1047.
- Bouslama L, Nasri D, Chollet L, Belguith K, Bourlet T, Aouni M, Pozzetto B, Pillet S. Natural recombination event within the capsid genomic region leading to a chimeric strain of human enterovirus B. *J Virol*. 2007 Sep;81(17):8944-52.
- Brandenburg B, Lee LY, Lakadamyali M, Rust MJ, Zhuang X, Hogle JM. Imaging poliovirus entry in live cells. *PLoS Biol*. 2007. 5(7):e183.
- Brown BA, Oberste MS, Alexander JP Jr, Kennett ML, Pallansch MA. Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. *J Virol*. 1999. 73(12):9969-75.
- Brown BA, Pallansch MA. Complete nucleotide sequence of enterovirus 71 is distinct from poliovirus. *Virus Res*. 1995. 39(2-3):195-205.
- Cai Y, Ku Z, Liu Q, Leng Q, Huang Z. A combination vaccine comprising of inactivated enterovirus 71 and coxsackievirus A16 elicits balanced protective immunity against both viruses. *Vaccine*. 2014 May. 32(21):2406-12.
- Cardoso MJ, Krishnan S, Tio PH, Perera D, Wong SC. Isolation of subgenus B adenovirus during a fatal outbreak of enterovirus 71-associated hand, foot, and mouth disease in Sibul, Sarawak. *Lancet*. 1999. 354(9183):987-91.
- Cardoso MJ, Perera D, Brown BA, Cheon D, Chan HM, Chan KP, Cho H, McMinn P. Molecular epidemiology of human enterovirus 71 strains and recent outbreaks in the Asia-Pacific region: comparative analysis of the VP1 and VP4 genes. *Emerg Infect Dis*. 2003. 9(4):461-8.
- Centers for Disease Control and Prevention (CDC). Severe hand, foot, and mouth disease associated with coxsackievirus A6 – Alabama, Connecticut, California, and Nevada, November 2011–February 2012. *MMWR*. 2012. 61:213–14
- Chakraborty R, Iturriza-Gómara M, Musoke R, Palakudy T, D'Agostino A, Gray J. An epidemic of enterovirus 71 infection among HIV-1-infected orphans in Nairobi. *AIDS*. 2004. 18(14):1968-70.
- Chambon M, Archimbaud C, Bailly JL, Gourgand JM, Charbonné F, Peigue-Lafeuille, H. Virucidal efficacy of glutaraldehyde against enteroviruses is related to the location of lysine



residues in exposed structures of the VP1 capsid protein. *Appl Environ Microbiol*. 2004 Mar;70(3):1717-22.

- Chan LG, Parashar UD, Lye MS, Ong FG, Zaki SR, Alexander JP, Ho KK, Han LL, Pallansch MA, Suleiman AB, Jegathesan M, Anderson LJ. Deaths of children during an outbreak of hand, foot, and mouth disease in sarawak, malaysia: clinical and pathological characteristics of the disease. For the Outbreak Study Group. *Clin Infect Dis*. 2000 Sep;31(3):678-83.
- Chan YF, AbuBakar S. Recombinant human enterovirus 71 in hand, foot and mouth disease patients. *Emerg Infect Dis*. 2004. 10(8):1468-70.
- Chang LY, King CC, Hsu KH, Ning HC, Tsao KC, Li CC, Huang YC, Shih SR, Chiou ST, Chen PY, Chang HJ, Lin TY. Risk factors of enterovirus 71 infection and associated hand, foot, and mouth disease/herpangina in children during an epidemic in Taiwan. *Pediatrics*. 2002. 109(6):e88.
- Chang LY, Lin TY, Huang YC, Tsao KC, Shih SR, Kuo ML, Ning HC, Chung PW, Kang CM. Comparison of enterovirus 71 and coxsackie-virus A16 clinical illnesses during the Taiwan enterovirus epidemic, 1998. *Pediatr Infect Dis J*. 1999. 18(12):1092-6.
- Chen LC, Yeh TM. Enterovirus 71 Infection of Human Immune Cells Induces the Production of Proinflammatory Cytokines. *J Biomed Lab Sc*. 2009. 21: 82–90.
- Chen P, Song Z, Qi Y, Feng X, Xu N, Sun Y, Wu X, Yao X, Mao Q, Li X, Dong W, Wan X, Huang N, Shen X, Liang Z, Li W. Molecular determinants of enterovirus 71 viral entry: cleft around GLN-172 on VP1 protein interacts with variable region on scavenger receptor B 2. *J Biol Chem*. 2012 Feb 24;287(9):6406-20.
- Chen X, Tan X, Li J, Jin Y, Gong L, Hong M, Shi Y, Zhu S, Zhang B, Zhang S, Zhang Y, Mao N, Xu W. Molecular epidemiology of coxsackievirus A16: intratype and prevalent intertype recombination identified. *PLoS One*. 2013. 8(12):e82861.
- Chen YC, Yu CK, Wang YF, Liu CC, Su IJ, Lei HY. A murine oral enterovirus 71 infection model with central nervous system involvement. *J Gen Virol*. 2004. 85(Pt 1):69-77.
- Chu PY, Lu PL, Tsai YL, Hsi E, Yao CY, Chen YH, Hsu LC, Wang SY, Wu HS, Lin YY, Su HJ, Lin KH. Spatiotemporal phylogenetic analysis and molecular characterization of coxsackievirus A4. *Infect Genet Evol*. 2011. 11(6):1426-35.
- Chumakov M, Voroshilova M, Shindarov L, Lavrova I, Gracheva L, Koroleva G, Vasilenko S, Brodvarova I, Nikolova M, Gyurova S, Gacheva M, Mitov G, Ninov N, Tsyilka E, Robinson I, Frolova M, Bashkirtsev V, Martiyanova L, Rodin V. Enterovirus 71 isolated from cases of epidemic poliomyelitis-like disease in Bulgaria. *Arch Virol*. 1979. 60(3-4):329-40.
- Chung SK, Kim JY, Kim IB, Park SI, Paek KH, Nam JH. Internalization and trafficking mechanisms of coxsackievirus B3 in HeLa cells. *Virology*. 2005. 333(1):31-40.
- Clark ME, Hämmerle T, Wimmer E, Dasgupta A. Poliovirus proteinase 3C converts an active form of transcription factor III C to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO J*. 1991. 10(10):2941-7.
- Clark ME, Lieberman PM, Berk AJ, Dasgupta A. Direct cleavage of human TATA-binding protein by poliovirus protease 3C in vivo and in vitro. *Mol Cell Biol*. 1993. 13(2):1232-7.
- Combelas N, Holmblat B, Joffret ML, Colbère-Garapin F, Delpeyroux F. Recombination between poliovirus and coxsackie A viruses of species C: a model of viral genetic plasticity and emergence. *Viruses*. 2011. 3(8):1460-84.
- Committee on the enteroviruses, National Foundation For Infantile Paralysis. 1957. *American Journal of public health*. 47(12):1556-66.
- Coyne CB, Bergelson JM. Virus-induced Abl and Fyn kinase signals permit coxsackievirus entry through epithelial tight junctions. *Cell*. 2006. 124(1):119-31.
- Coyne CB, Kim KS, Bergelson JM. Poliovirus entry into human brain microvascular cells requires receptor-induced activation of SHP-2. *EMBO J*. 2007 Sep 5;26(17):4016-28.
- Dang M, Wang X, Wang Q, Wang Y, Lin J, Sun Y, Li X, Zhang L, Lou Z, Wang J, Rao Z. Molecular mechanism of SCARB2-mediated attachment and uncoating of EV71. *Protein Cell*. 2014 Sep;5(9):692-703.

- Dang S, Gao N, Li Y, Li M, Wang X, Jia X, Zhai S, Zhang X, Liu J, Deng H, Dong T. Dominant CD4-dependent RNA-dependent RNA polymerase-specific T-cell responses in children acutely infected with human enterovirus 71 and healthy adult controls. *Immunology*. 2014. 142(1):89-100.
- de Breyne S, Bonderoff JM, Chumakov KM, Lloyd RE, Hellen CU. Cleavage of eukaryotic initiation factor eIF5B by enterovirus 3C proteases. *Virology*. 2008. 378(1):118-22.
- Deibel R, Gross LL, Collins DN. Isolation of a new enterovirus (38506). *Proc Soc Exp Biol Med*. 1975. 148(1):203-7.
- Delsuc F. & Douzery E. J. P. *Les méthodes probabilistes en phylogénie moléculaire. (1) Modèles d'évolution des séquences et maximum de vraisemblance*. Biosystema : « Avenir et pertinence des méthodes d'analyse en phylogénie moléculaire ». 2004. 22 : 59-74.
- Deshpande JM, Nadkarni SS, Francis PP. Enterovirus 71 isolated from a case of acute flaccid paralysis in India represents a new genotype. *Current Science*. 2003. 84: 1350–1353.
- Diedrich S, Weinbrecht A, Schreier E. Seroprevalence and molecular epidemiology of enterovirus 71 in Germany. *Arch Virol*. 2009. 154(7):1139-42.
- Doedens JR, Kirkegaard K. Inhibition of cellular protein secretion by poliovirus proteins 2B and 3A. *EMBO J*. 1995 Mar 1;14(5):894-907.
- Dotzauer A, Kraemer L. Innate and adaptive immune responses against picornaviruses and their counteractions: An overview. *World J Virol*. 2012. 1(3):91-107.
- Douzery E. J. P., Delsuc F. & Philippe H. *Les datations moléculaires à l'heure de la génomique*. Médecine Sciences. 2006. 22 (4) 374-380.
- Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. *PLoS Biol*. 2006. 4(5):e88.
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol*. 2005. 22(5):1185-92.
- Emerson B, Paradis CE & Thebaud C. Revealing the demographic histories of species with DNA sequences. *Trends in Ecology and Evolution*. 2001. 16:707-716.
- Excoffier L, Heckel G. Computer programs for population genetics data analysis: a survival guide. *Nat Rev Genet*. 2006. 7(10):745-58.
- Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, Tatem AJ, Sousa JD, Arinaminpathy N, Pèpin J, Posada D, Peeters M, Pybus OG, Lemey P. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014. 346(6205):56-61.
- Felsenstein J. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet*. 1973. 25(5):471-92.
- Feng Q, Langereis MA, Lork M, Nguyen M, Hato SV, Lanke K, Emdad L, Bhoopathi P, Fisher PB, Lloyd RE, van Kuppeveld FJ. Enterovirus 2Apro targets MDA5 and MAVS in infected cells. *J Virol*. 2014. 88(6):3369-78.
- Fischer TK, Nielsen AY, Sydenham TV, Andersen PH, Andersen B, Midgley SE. Emergence of enterovirus 71 C4a in Denmark, 2009 to 2013. *Euro Surveill*. 2014. 19(38).
- Fogg MH, Teterina NL, Ehrenfeld E. Membrane requirements for uridylylation of the poliovirus VPg protein and viral RNA synthesis in vitro. *J Virol*. 2003. 77(21):11408-16.
- Fonseca MC, Sarmiento L, Resik S, Martínez Y, Hung LH, Morier L, Piñón A, Valdéz O, Kourí V, González G. Coxsackievirus A6 and enterovirus 71 causing hand, foot and mouth disease in Cuba, 2011-2013. *Arch Virol*. 2014 Apr 10.
- Freistadt MS, Vaccaro JA, Eberle KE. Biochemical characterization of the fidelity of poliovirus RNA-dependent RNA polymerase. *Virol J*. 2007. 24;4:44.
- Fujimoto T, Chikahira M, Yoshida S, Ebira H, Hasegawa A, Totsuka A, Nishio O. Outbreak of central nervous system disease associated with hand, foot, and mouth disease in Japan during the summer of 2000: detection and molecular epidemiology of enterovirus 71. *Microbiol Immunol*. 2002. 46(9):621-7.

- Gamarnik AV, Andino R. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* 1998. 12(15):2293-304.
- Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol.* 1997. 14(7):685-95.
- Gibbs MJ, Armstrong JS, Gibbs AJ. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics.* 2000. 16(7): 573–582.
- Gilbert GL, Dickson KE, Waters MJ, Kennett ML, Land SA, Sneddon M. Outbreak of enterovirus 71 infection in Victoria, Australia, with a high incidence of neurologic involvement. *Pediatr Infect Dis J.* 1988. 7(7):484-8.
- Gmyl AP, Belousov EV, Maslova SV, Khitrina EV, Chetverin AB, Agol VI. Nonreplicative RNA recombination in poliovirus. *J Virol.* 1999. 73(11):8958-65.
- Gmyl AP, Korshenko SA, Belousov EV, Khitrina EV, Agol VI. Nonreplicative homologous RNA recombination: promiscuous joining of RNA pieces? *RNA.* 2003. 9(10):1221-31.
- Gojobori T, Moriyama EN, Kimura M. Molecular clock of viral evolution, and the neutral theory. *Proc Natl Acad Sci U S A.* 1990. 87(24):10015-8.
- Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 2010. 27(2):221-4.
- Gradi A, Imataka H, Svitkin YV, Rom E, Raught B, Morino S, Sonenberg N. A novel functional human eukaryotic translation initiation factor 4G. *Mol Cell Biol.* 1998. 18(1):334-42.
- Gradi A, Svitkin YV, Imataka H, Sonenberg N. Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc Natl Acad Sci U S A.* 1998. 95(19):11089-94.
- Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, Holmes EC. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science.* 2004 Jan 16;303(5656):327-32.
- Greve JM, Davis G, Meyer AM, Forte CP, Yost SC, Marlor CW, Kamarck ME, McClelland A. The major human rhinovirus receptor is ICAM-1. *Cell.* 1989 Mar 10;56(5):839-47.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 2003. 52(5):696-704.
- Gullberg M, Tolf C, Jonsson N, Mulders MN, Savolainen-Kopra C, Hovi T, Van Ranst M, Lemey P, Hafenstein S, Lindberg AM. Characterization of a putative ancestor of coxsackievirus B5. *J Virol.* 2010. 84(19):9695-708.
- Han JF, Xu S, Zhang Y, Zhu SY, Wu DL, Yang XD, Liu H, Sun BX, Wu XY, Qin CF. Hand, foot, and mouth disease outbreak caused by coxsackievirus A6, China, 2013. *J Infect.* 2014. 69(3):303-5.
- Hashimoto I, Hagiwara A, Kodama H. Neurovirulence in cynomolgus monkeys of enterovirus 71 isolated from a patient with hand, foot and mouth disease. *Arch Virol.* 1978. 56(3):257-61.
- Hastings W K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika.* 1970. 57(1):97-109
- Hayman R, Shepherd M, Tarring C, Best E. Outbreak of variant hand-foot-and-mouth disease caused by coxsackievirus A6 in Auckland, New Zealand. *J Paediatr Child Health.* 2014 Aug 15. doi: 10.1111/jpc.12708.
- He YQ, Chen L, Xu WB, Yang H, Wang HZ, Zong WP, Xian HX, Chen HL, Yao XJ, Hu ZL, Luo M, Zhang HL, Ma HW, Cheng JQ, Feng QJ, Zhao DJ. Emergence, circulation, and spatiotemporal phylogenetic analysis of coxsackievirus a6- and coxsackievirus a10-associated hand, foot, and mouth disease infections from 2008 to 2012 in Shenzhen, China. *J Clin Microbiol.* 2013. 51(11):3560-6.
- Heeger PS, Lalli PN, Lin F, Valujskikh A, Liu J, Muqim N, Xu Y, Medof ME. Decay-accelerating factor modulates induction of T cell immunity. *J Exp Med.* 2005 May 16;201(10):1523-30.
- Heikkilä O, Susi P, Stanway G, Hyypiä T. Integrin alphaVbeta6 is a high-affinity receptor for coxsackievirus A9. *J Gen Virol.* 2009. 90(Pt 1):197-204.

- Henquell C, Mirand A, Richter J, Schuffenecker I, Böttiger B, Diedrich S, Terletskaia-Ladwig E, Christodoulou C, Peigue-Lafeuille H, Bailly JL. Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission. *J Virol*. 2013. 87(22):12249-59.
- Ho M, Chen ER, Hsu KH, Twu SJ, Chen KT, Tsai SF, Wang JR, Shih SR. An epidemic of enterovirus 71 infection in Taiwan. Taiwan Enterovirus Epidemic Working Group. *N Engl J Med*. 1999. 341(13):929-35.
- Horsley E, Just E, Torres C, Huhtinen E, Forssman B, Slade R. Enterovirus 71 outbreak in Northern Sydney, 2013: Case series and initial response. *J Paediatr Child Health*. 2014. 50(7):525-30.
- Hosoya M, Kawasaki Y, Sato M, Honzumi K, Hayashi A, Hiroshima T, Ishiko H, Kato K, Suzuki H. Genetic diversity of coxsackievirus A16 associated with hand, foot, and mouth disease epidemics in Japan from 1983 to 2003. *J Clin Microbiol*. 2007. 45(1):112-20.
- Hosoya M, Kawasaki Y, Sato M, Honzumi K, Hayashi A, Hiroshima T, Ishiko H, Kato K, Suzuki H. Genetic diversity of coxsackievirus A16 associated with hand, foot, and mouth disease epidemics in Japan from 1983 to 2003. *J Clin Microbiol*. 2007. 45(1):112-20.
- Hu YF, Yang F, Du J, Dong J, Zhang T, Wu ZQ, Xue Y, Jin Q. Complete genome analysis of coxsackievirus A2, A4, A5, and A10 strains isolated from hand, foot, and mouth disease patients in China revealing frequent recombination of human enterovirus A. *J Clin Microbiol*. 2011. 49(7):2426-34.
- Huang ML, Chiang PS, Chia MY, Luo ST, Chang LY, Lin TY, Ho MS, Lee MS. Cross-reactive neutralizing antibody responses to enterovirus 71 infections in young children: implications for vaccine development. *PLoS Negl Trop Dis*. 2013. 7(2):e2067.
- Huang SW, Hsu YW, Smith DJ, Kiang D, Tsai HP, Lin KH, Wang SM, Liu CC, Su IJ, Wang JR. Reemergence of enterovirus 71 in 2008 in taiwan: dynamics of genetic and antigenic evolution from 1998 to 2008. *J Clin Microbiol*. 2009. 47(11):3653-62.
- Huang YP, Lin TL, Kuo CY, Lin MW, Yao CY, Liao HW, Hsu LC, Yang CF, Yang JY, Chen PJ, Wu HS. The circulation of subgenogroups B5 and C5 of enterovirus 71 in Taiwan from 2006 to 2007. *Virus Res*. 2008. 137(2):206-12.
- Huemer HP, Ortner B, Huang CW, Schmid D, Mutz I, Wewalka G, Yang JY, Allerberger F. Isolating Asian enterovirus 71 subgenogroup C4 in two Austrian clinical samples from 2004. *Euro Surveill*. 2008 13(28).
- Ibrahim W, Boukhadra N, Nasri-Zoghalmi D, Berthelot P, Omar S, Bourlet T, Pozzetto B, Pillet S. Partial sequencing of the VP2 capsid gene for direct enterovirus genotyping in clinical specimens. *Clin Microbiol Infect*. 2014 Sep;20(9):O558-65.
- Ishimaru Y, Nakano S, Yamaoka K, Takami S. Outbreaks of hand, foot, and mouth disease by enterovirus 71. High incidence of complication disorders of central nervous system. *Arch Dis Child*. 1980. 55(8):583-8.
- Iwai M, Masaki A, Hasegawa S, Obara M, Horimoto E, Nakamura K, Tanaka Y, Endo K, Tanaka K, Ueda J, Shiraki K, Kurata T, Takizawa T. Genetic changes of coxsackievirus A16 and enterovirus 71 isolated from hand, foot, and mouth disease patients in Toyama, Japan between 1981 and 2007. *Jpn J Infect Dis*. 2009. 62(4):254-9.
- Jackson WT, Giddings TH Jr, Taylor MP, Mulinyawe S, Rabinovitch M, Kopito RR, Kirkegaard K. Subversion of cellular autophagosomal machinery by RNA viruses. *PLoS Biol*. 2005. 3(5):e156.
- Jacobson MF, Baltimore D. Polypeptide cleavages in the formation of poliovirus proteins. *Proc Natl Acad Sci U S A*. 1968. 61(1):77-84.
- Jee YM, Cheon DS, Kim K, Cho JH, Chung YS, Lee J, Lee SH, Park KS, Lee JH, Kim EC, Chung HJ, Kim DS, Yoon JD, Cho HW. Genetic analysis of the VP1 region of human enterovirus 71 strains isolated in Korea during 2000. *Arch Virol*. 2003. 148(9):1735-46.
- Jenkins GM, Rambaut A, Pybus OG, Holmes EC. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol*. 2002. 54(2):156-65.

- Joachims M, Van Breugel PC, Lloyd RE. Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol*. 1999. 73(1):718-27.
- Kapusinszky B, Szomor KN, Farkas A, Takács M, Berencsi G. Detection of non-polio enteroviruses in Hungary 2000-2008 and molecular epidemiology of Enterovirus 71, coxsackievirus A16, and echovirus 30. *Virus Genes*. 2010. 40(2):163-73.
- Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, Uematsu S, Jung A, Kawai T, Ishii KJ, Yamaguchi O, Otsu K, Tsujimura T, Koh CS, Reis e Sousa C, Matsuura Y, Fujita T, Akira S. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature*. 2006 May 4;441(7089):101-5.
- Kehle J, Roth B, Metzger C, Pfitzner A, Enders G. Molecular characterization of an Enterovirus 71 causing neurological disease in Germany. *J Neurovirol*. 2003. 9(1):126-8.
- Kennett ML, Birch CJ, Lewis FA, Yung AP, Locarnini SA, Gust ID. Enterovirus type 71 infection in Melbourne. *Bull World Health Organ*. 1974. 51(6):609-15.
- Khanh TH, Sabanathan S, Thanh TT, Thoa le PK, Thuong TC, Hang Vt, Farrar J, Hien TT, Chau Nv, van Doorn HR. Enterovirus 71-associated hand, foot, and mouth disease, Southern Vietnam, 2011. *Emerg Infect Dis*. 2012. 18(12):2002-5.
- Khong WX, Yan B, Yeo H, Tan EL, Lee JJ, Ng JK, Chow VT, Alonso S. A non-mouse-adapted enterovirus 71 (EV71) strain exhibits neurotropism, causing neurological manifestations in a novel mouse model of EV71 infection. *J Virol*. 2012. 86(4):2121-31.
- King AM, McCahon D, Slade WR, Newman JW. Recombination in RNA. *Cell*. 1982. 29(3):921-8.
- Kingman J.F.C. The coalescent. *Stochastic Processes and their Applications*. 13(3):235–248
- Kirkegaard K, Baltimore D. The mechanism of RNA recombination in poliovirus. *Cell*. 1986. 47(3):433-43.
- Kobayashi M, Makino T, Hanaoka N, Shimizu H, Enomoto M, Okabe N, Kanou K, Konagaya M, Oishi K, Fujimoto T. Clinical manifestations of coxsackievirus A6 infection associated with a major outbreak of hand, foot, and mouth disease in Japan. *Jpn J Infect Dis*. 2013. 66(3):260-1.
- Koike S, Ise I, Nomoto A. Functional domains of the poliovirus receptor. *Proc Natl Acad Sci U S A*. 1991 May 15;88(10):4104-8.
- Kondrashova A, Hyöty H. Role of viruses and other microbes in the pathogenesis of type 1 diabetes. *Int Rev Immunol*. 2014. 33(4):284-95.
- Koroleva GA, Karmysheva VY, Lukashev AN. Enterovirus 71 pathogenicity in monkeys and cotton rats. *Arch Virol*. 2014. 159(5):1133-8.
- Kung SH, Wang SF, Huang CW, Hsu CC, Liu HF, Yang JY. Genetic and antigenic analyses of enterovirus 71 isolates in Taiwan during 1998-2005. *Clin Microbiol Infect*. 2007. 13(8):782-7.
- Kung YA, Hung CT, Liu YC, Shih SR. Update on the development of enterovirus 71 vaccines. *Expert Opin Biol Ther*. 2014 Oct;14(10):1455-64.
- Lam TT, Hon CC, Lemey P, Pybus OG, Shi M, Tun HM, Li J, Jiang J, Holmes EC, Leung FC. Phylodynamics of H5N1 avian influenza virus in Indonesia. *Mol Ecol*. 2012 Jun;21(12):3062-77.
- Ledinko N. Genetic recombination with poliovirus type 1. Studies of crosses between a normal horse serum-resistant mutant and several guanidine-resistant mutants of the same strain. *Virology*. 1963. 20:107-19.
- Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. *PLoS Comput Biol*. 2009. 5(9):e1000520.
- Li JX, Mao QY, Liang ZL, Ji H, Zhu FC. Development of enterovirus 71 vaccines: from the lab bench to Phase III clinical trials. *Expert Rev Vaccines*. 2014 May;13(5):609-18.
- Li L, He Y, Yang H, Zhu J, Xu X, Dong J, Zhu Y, Jin Q. Genetic characteristics of human enterovirus 71 and coxsackievirus A16 circulating from 1999 to 2004 in Shenzhen, People's Republic of China. *J Clin Microbiol*. 2005. 43(8):3835-9.

- Liang CC, Sun MJ, Lei HY, Chen SH, Yu CK, Liu CC, Wang JR, Yeh TM. Human endothelial cell activation and apoptosis induced by enterovirus 71 infection. *J Med Virol.* 2004 Dec;74(4):597-603.
- Liang ZL, Mao QY, Wang YP, Zhu FC, Li JX, Yao X, Gao F, Wu X, Xu M, Wang JZ. Progress on the research and development of inactivated EV71 whole-virus vaccines. *Hum Vaccin Immunother.* 2013. 9(8):1701-5.
- Lim, K. A., Benyesh-Melnick, M. 1960. Typing of viruses by combinations of antiserum pools. Application to typing of Enteroviruses (Coxsackie and ECHO). *J Immunol.* 84:309-317.
- Lin YW, Wang SW, Tung YY, Chen SH. Enterovirus 71 infection of human dendritic cells. *Exp Biol Med (Maywood).* 2009. 234(10):1166-73.
- Lindberg AM, Andersson P, Savolainen C, Mulders MN, Hovi T. Evolution of the genome of Human enterovirus B: incongruence between phylogenies of the VP1 and 3CD regions indicates frequent recombination within the species. *J Gen Virol.* 2003. 84(Pt 5):1223-35.
- Linsuwanon P, Puenpa J, Huang SW, Wang YF, Mauleekoonphairoj J, Wang JR, Poovorawan Y. Epidemiology and seroepidemiology of human enterovirus 71 among Thai populations. *J Biomed Sci.* 2014. 21:16.
- Lu J, Yi L, Zhao J, Yu J, Chen Y, Lin MC, Kung HF, He ML. Enterovirus 71 disrupts interferon signaling by reducing the level of interferon receptor 1. *J Virol.* 2012. 86(7):3767-76.
- Luchs A, Cilli A, Russo DH, Costa FF, Carmona Rde C, Timenetsky Mdo C. Evaluation of enterovirus 71 immune status in São Paulo state, Brazil. *Rev Inst Med Trop Sao Paulo.* 2010. 52(6):339-41.
- Lukashev AN, Shumilina EY, Belalov IS, Ivanova OE, Eremeeva TP, Reznik VI, Trotsenko OE, Drexler JF, Drosten C. Recombination strategies and evolutionary dynamics of the Human enterovirus A global gene pool. *J Gen Virol.* 2014. 95(Pt4):868-73.
- Lum LC, Wong KT, Lam SK, Chua KB, Goh AY, Lim WL, Ong BB, Paul G, AbuBakar S, Lambert M. Fatal enterovirus 71 encephalomyelitis. *J Pediatr.* 1998. 133(6):795-8.
- Martin DP, Posada D, Crandall K, Williamson C (2005) A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Res Hum Retroviruses* 21(1): 98–102.
- Martin DP, Rybicki E (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16(16): 562–563.
- Mathes EF, Oza V, Frieden IJ, Cordero KM, Yagi S, Howard R, Kristal L, Ginocchio CC, Schaffer J, Maguiness S, Bayliss S, Lara-Corrales I, Garcia-Romero MT, Kelly D, Salas M, Oberste MS, Nix WA, Glaser C, Antaya R. "Eczema coxsackium" and unusual cutaneous findings in an enterovirus outbreak. *Pediatrics.* 2013. 132(1):e149-57.
- McMinn P, Stratov I, Nagarajan L, Davis S. Neurological manifestations of enterovirus 71 infection in children during an outbreak of hand, foot, and mouth disease in Western Australia. *Clin Infect Dis.* 2001. 32(2):236-42.
- McMinn PC. An overview of the evolution of enterovirus 71 and its clinical and public health significance. *FEMS Microbiol Rev.* 2002. 26(1):91-107.
- McMinn PC. Recent advances in the molecular epidemiology and control of human enterovirus 71 infection. *Curr Opin Virol.* 2012. 2(2):199-205.
- McWilliam Leitch EC, Cabrerizo M, Cardoso J, Harvala H, Ivanova OE, Kroes AC, Lukashev A, Muir P, Odoom J, Roivainen M, Susi P, Trallero G, Evans DJ, Simmonds P. Evolutionary dynamics and temporal/geographical correlates of recombination in the human enterovirus echovirus types 9, 11, and 30. *J Virol.* 2010 Sep;84(18):9292-300.
- McWilliam Leitch EC, Cabrerizo M, Cardoso J, Harvala H, Ivanova OE, Koike S, Kroes AC, Lukashev A, Perera D, Roivainen M, Susi P, Trallero G, Evans DJ, Simmonds P. The association of recombination events in the founding and emergence of subgenogroup evolutionary lineages of human enterovirus 71. *J Virol.* 2012. 86(5):2676-85.
- Melnick JL. Enterovirus type 71 infections: a varied clinical pattern sometimes mimicking paralytic poliomyelitis. *Rev Infect Dis.* 1984. 6 Suppl 2:S387-90.

- Melnick JL. Thermostability of poliovirus and measles vaccines. *Dev Biol Stand.* 1996;87:155-60. Review.
- Mendelsohn CL, Wimmer E, Racaniello VR. Cellular receptor for poliovirus: olecular cloning, nucleotide sequence, and expression of a new member of the mmunoglobulin superfamily. *Cell.* 1989 Mar 10;56(5):855-65.
- Minor PD, Ferguson M, Evans DM, Almond JW, Icenogle JP. Antigenic structure of polioviruses of serotypes 1, 2 and 3. *J Gen Virol.* 1986 Jul;67 ( Pt 7):1283-91.
- Mirand A, Henquell C, Archimbaud C, Chambon M, Charbonné F, Peigue-Lafeuille H, Bailly JL. Prospective identification of enteroviruses involved in meningitis in 2006 through direct genotyping in cerebrospinal fluid. *J ClinMicrobiol.* 2008. 46 87–96.
- Mirand A, Henquell C, Archimbaud C, Ughetto S, Antona D, Bailly JL, Peigue-Lafeuille H. Outbreak of hand, foot and mouth disease/herpangina associated with coxsackievirus A6 and A10 infections in 2010, France: a large citywide, prospective observational study. *Clin Microbiol Infect.* 2012. 18(5):E110-8.
- Mirand A, Schuffenecker I, Henquell C, Billaud G, Jugie G, Falcon D, Mahul A, Archimbaud C, Terletskaia-Ladwig E, Diedrich S, Huemer HP, Enders M, Lina B, Peigue-Lafeuille H, Bailly JL. Phylogenetic evidence for a recent spread of two populations of human enterovirus 71 in European countries. *J Gen Virol.* 2010 Sep;91(Pt 9):2263-77.
- Miyamura K, Nishimura Y, Abo M, Wakita T, Shimizu H. Adaptive mutations in the genomes of enterovirus 71 strains following infection of mouse cells expressing human P-selectin glycoprotein ligand-1. *J Gen Virol.* 2011 Feb;92(Pt 2):287-91.
- Mizuta K, Abiko C, Aoki Y, Ikeda T, Matsuzaki Y, Hongo S, Itagaki T, Katsushima N, Ohmi A, Nishimura H, Ahiko T. Molecular epidemiology of Coxsackievirus A16 strains isolated from children in Yamagata, Japan between 1988 and 2011. *Microbiol Immunol.* 2013. 57(5):400-5.
- Mizuta K, Abiko C, Murata T, Matsuzaki Y, Itagaki T, Sanjoh K, Sakamoto M, Hongo S, Murayama S, Hayasaka K. Frequent importation of enterovirus 71 from surrounding countries into the local community of Yamagata, Japan, between 1998 and 2003. *J Clin Microbiol.* 2005. 43(12):6171-5.
- Mizuta K, Aoki Y, Suto A, Ootani K, Katsushima N, Itagaki T, Ohmi A, Okamoto M, Nishimura H, Matsuzaki Y, Hongo S, Sugawara K, Shimizu H, Ahiko T. Cross-antigenicity among EV71 strains from different genogroups isolated in Yamagata, Japan, between 1990 and 2007. *Vaccine.* 2009 May. 27(24):3153-8.
- Montes M, Artieda J, Piñeiro LD, Gastesi M, Diez-Nieves I, Cilla G. Hand, foot, and mouth disease outbreak and coxsackievirus A6, northern Spain, 2011. *Emerg Infect Dis.* 2013. 19(4).
- Nagy G, Takátsy S, Kukán E, Mihály I, Dömök I. Virological diagnosis of enterovirus type 71 infections: experiences gained during an epidemic of acute CNS diseases in Hungary in 1978. *Arch Virol.* 1982. 71(3):217-27.
- Nishimura Y, Lee H, Hafenstein S, Kataoka C, Wakita T, Bergelson JM, Shimizu H. Enterovirus 71 binding to PSGL-1 on leukocytes: VP1-145 acts as a molecular switch to control receptor interaction. *PLoS Pathog.* 2013;9(7):e1003511.
- Nishimura Y, Shimizu H. Cellular receptors for human enterovirus species A. *Front Microbiol.* 2012 Mar 27;3:105.
- Nishimura Y, Shimojima M, Tano Y, Miyamura T, Wakita T, Shimizu H. Human P-selectin glycoprotein ligand-1 is a functional receptor for enterovirus 71. *Nat Med.* 2009 Jul;15(7):794-7.
- Nishimura Y, Wakita T, Shimizu H. Tyrosine sulfation of the amino terminus of PSGL-1 is critical for enterovirus 71 infection. *PLoS Pathog.* 2010 Nov 4;6(11):e1001174.
- Nix WA, Jiang B, Maher K, Strobert E, Oberste MS. Identification of enteroviruses in naturally infected captive primates. *J Clin Microbiol.* 2008. 46(9):2874-8.
- Nugent CI, Johnson KL, Sarnow P, Kirkegaard K. Functional coupling between replication and packaging of poliovirus replicon RNA. *J Virol.* 1999. 73(1):427-35.

- Oberste MS, Maher K, Kilpatrick DR, Pallansch MA. Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. *J Virol*. 1999. 73(3):1941-8.
- Oberste MS, Peñaranda S, Maher K, Pallansch MA. Complete genome sequences of all members of the species Human enterovirus A. *J Gen Virol*. 2004 Jun;85(Pt 6):1597-607. PubMed PMID: 15166444. Ogra PL. Mucosal immune response to poliovirus vaccines in childhood. *Rev Infect Dis*. 1984 May-Jun;6 Suppl 2:S361-8.
- Ooi MH, Solomon T, Podin Y, Mohan A, Akin W, Yusuf MA, del Sel S, Kontol KM, Lai BF, Clear D, Chieng CH, Blake E, Perera D, Wong SC, Cardoso J. Evaluation of different clinical sample types in diagnosis of human enterovirus 71-associated hand-foot-and-mouth disease. *J Clin Microbiol*. 2007. 45(6):1858-66.
- Ooi MH, Wong SC, Lewthwaite P, Cardoso MJ, Solomon T. Clinical features, diagnosis, and management of enterovirus 71. *Lancet Neurol*. 2010. 9(11):1097-105.
- Padidam M, Sawyer S, Fauquet CM (1999) Possible emergence of new geminiviruses by frequent recombination. *Virology* 265(2): 218–225.
- Pallansch M. and Roos R. 2007. Enteroviruses: Polioviruses, Coxsackieviruses, and Newer Enteroviruses. In *Fields Virology*, 5th Ed Knipe D. M., Howley P. M., eds. Lippincott Williams & Wilkins, Philadelphia, p. 840–893.
- Palmenberg AC. In vitro synthesis and assembly of picornaviral capsid intermediate structures. *J Virol*. 1982. 44(3):900-6.
- Park N, Katikaneni P, Skern T, Gustin KE. Differential targeting of nuclear pore complex proteins in poliovirus-infected cells. *J Virol*. 2008. 82(4):1647-55.
- Peigue-Lafeuille H, Mirand A, Archimbaud C, Bailly J-L, Henquell C. Émergence et réémergence chez les entérovirus : de la poliomyélite à la maladie pieds-mains-bouche. *Virologie*. 2014. 18 (2):87-104.
- Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*. 1988. 334(6180):320-5.
- Perera D, Yusof MA, Podin Y, Ooi MH, Thao NT, Wong KK, Zaki A, Chua KB, Malik YA, Tu PV, Tien NT, Puthavathana P, McMinn PC, Cardoso MJ. Molecular phylogeny of modern coxsackievirus A16. *Arch Virol*. 2007. 152(6):1201-8.
- Pérez-Vélez CM, Anderson MS, Robinson CC, McFarland EJ, Nix WA, Pallansch MA, Oberste MS, Glodé MP. Outbreak of neurologic enterovirus type 71 disease: a diagnostic challenge. *Clin Infect Dis*. 2007. 45(8):950-7.
- Pilipenko EV, Gmyl AP, Agol VI. A model for rearrangements in RNA genomes. *Nucleic Acids Res*. 1995. 23(11):1870-5.
- Podin Y, Gias EL, Ong F, Leong YW, Yee SF, Yusof MA, Perera D, Teo B, Wee TY, Yao SC, Yao SK, Kiyu A, Arif MT, Cardoso MJ. Sentinel surveillance for human enterovirus 71 in Sarawak, Malaysia: lessons from the first 7 years. *BMC Public Health*. 2006. 6:180.
- Polio vaccines and polio immunization in the pre-eradication era: WHO position paper. *Wkly Epidemiol Rec*. 2010 Jun 4;85(23):213-28.
- Posada D, Crandall KA (2001) Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc Natl Acad Sci* 98(24): 13757–13762.
- Puenpa J, Chieochansin T, Linsuwanon P, Korkong S, Thongkomplew S, Vichaiwattana P, Theamboonlers A, Poovorawan Y. Hand, foot, and mouth disease caused by coxsackievirus A6, Thailand, 2012. *Emerg Infect Dis*. 2013. 19(4):641-3.
- Pybus OG, Rambaut A, Harvey PH. An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*. 2000. 155(3):1429-37.
- Pybus OG, Rambaut A. GENIE: estimating demographic history from molecular phylogenies. *Bioinformatics*. 2002. 18(10):1404-5.
- Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet*. 2009.10(8):540-50.



- Rabenau HF, Richter M, Doerr HW. Hand, foot and mouth disease: seroprevalence of Coxsackie A16 and Enterovirus 71 in Germany. *Med Microbiol Immunol*. 2010. 199(1):45-51.
- Racaniello, V. R. 2007. Picornaviridae: the viruses and their replication. In *Fields Virology*, 5th Ed Knipe D. M., Howley P. M., eds. Lippincott Williams & Wilkins, Philadelphia, p. 796–839.
- Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. *Nature*. 2008. 453(7195):615-9.
- Rannala B, Yang Z. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J Mol Evol*. 1996. 43(3):304-11.
- Ren XX, Ma L, Liu QW, Li C, Huang Z, Wu L, Xiong SD, Wang JH, Wang HB. The molecule of DC-SIGN captures enterovirus 71 and confers dendritic cell-mediated viral trans-infection. *Virology*. 2014. 11:47.
- Richter J, Koptides D, Tryfonos C, Christodoulou C. Molecular typing of enteroviruses associated with viral meningitis in Cyprus, 2000-2002. *J Med Microbiol*. 2006. 55(Pt 8):1035-41.
- Roivainen M, Piirainen L, Hovi T, Virtanen I, Riikonen T, Heino J, Hyypiä T. Entry of coxsackievirus A9 into host cells: specific interactions with alpha v beta 3 integrin, the vitronectin receptor. *Virology*. 1994. 203(2):357-65.
- Rosenberg NA, Nordborg M. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat Rev Genet*. 2002. 3(5):380-90.
- Rueckert RR, Dunker AK, Stoltzfus CM. The structure of mouse-Elberfeld virus: a model. *Proc Natl Acad Sci U S A*. 1969 Mar;62(3):912-9.
- Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987. 4(4):406-25.
- Samuda GM, Chang WK, Yeung CY, Tang PS. Monoplegia caused by Enterovirus 71: an outbreak in Hong Kong. *Pediatr Infect Dis J*. 1987. 6(2):206-8.
- Saxena VK, Sane S, Nadkarni SS, Sharma DK, Deshpande JM. Genetic diversity of enterovirus A71, India. *Emerg Infect Dis*. 2015. 21(1):123-6.
- Schlegel A, Giddings TH Jr, Ladinsky MS, Kirkegaard K. Cellular origin and ultrastructure of membranes induced during poliovirus infection. *J Virol*. 1996. 70(10):6576-88.
- Schmidt NJ, Ho HH, Lennette EH. Propagation and isolation of group A coxsackieviruses in RD cells. *J Clin Microbiol*. 1975. 2(3):183-5.
- Schmidt NJ, Lennette EH, Ho HH. An apparently new enterovirus isolated from patients with disease of the central nervous system. *J Infect Dis*. 1974. 129(3):304-9.
- Schuffenecker I, Henquell C, Mirand A, Coste-Burel M, Marque-Juillet S, Desbois D, Lagathu G, Bornebusch L, Bailly JL, Lina B. New introductions of enterovirus 71 subgenogroup c4 strains, France, 2012. *Emerg Infect Dis*. 2014. 20(8):1343-6.
- Schuffenecker I, Mirand A, Antona D, Henquell C, Chomel JJ, Archimbaud C, Billaud G, Peigue-Lafeuille H, Lina B, Bailly JL. Epidemiology of human enterovirus 71 infections in France, 2000-2009. *J Clin Virol*. 2011. 50(1):50-6.
- Shafren DR, Dorahy DJ, Greive SJ, Burns GF, Barry RD. Mouse cells expressing human intercellular adhesion molecule-1 are susceptible to infection by coxsackievirus A21. *J Virol*. 1997 Jan;71(1):785-9.
- Shafren DR, Dorahy DJ, Ingham RA, Burns GF, Barry RD. Coxsackievirus A21 binds to decay-accelerating factor but requires intercellular adhesion molecule 1 for cell entry. *J Virol*. 1997 Jun;71(6):4736-43.
- Shieh JT, Bergelson JM. Interaction with decay-accelerating factor facilitates coxsackievirus B infection of polarized epithelial cells. *J Virol*. 2002 Sep;76(18):9474-80.
- Shulman LM, Gavrilin E, Jorba J, Martin J, Burns CC, Manor Y, Moran-Gilad J, Sofer D, Hindiyeh MY, Gamzu R, Mendelson E, Grotto I; Genotype – Phenotype Identification (GPI) group. Molecular epidemiology of silent introduction and sustained transmission of wild poliovirus type 1, Israel, 2013. *Euro Surveill*. 2014. 19(7):20709.

- Simmonds P, Welch J. Frequency and dynamics of recombination within different species of human enteroviruses. *J Virol*. 2006. 80(1):483-93.
- Sinclair C, Gaunt E, Simmonds P, Broomfield D, Nwafor N, Wellington L, Templeton K, Willocks L, Schofield O, Harvala H. Atypical hand, foot, and mouth disease associated with coxsackievirus A6 infection, Edinburgh, United Kingdom, January to February 2014. *Euro Surveill*. 2014. 19(12):20745.
- Singh S, Poh CL, Chow VT. Complete sequence analyses of enterovirus 71 strains from fatal and non-fatal cases of the hand, foot and mouth disease outbreak in Singapore (2000). *Microbiol Immunol*. 2002. 46(11):801-8.
- Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghvani J, Bhatt S, Peiris JS, Guan Y, Rambaut A. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*. 2009. 459(7250):1122-5.
- Smith JM Analyzing the mosaic structure of genes. *J Mol Evol*. 1992. 34(2): 126–129.
- Sobo K, Rubbia-Brandt L, Brown TD, Stuart AD, McKee TA. Decay-accelerating factor binding determines the entry route of echovirus 11 in polarized epithelial cells. *J Virol*. 2011 Dec;85(23):12376-86.
- Solomon T, Lewthwaite P, Perera D, Cardoso MJ, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect Dis*. 2010. 10(11):778-90.
- Suhy DA, Giddings TH Jr, Kirkegaard K. Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. *J Virol*. 2000. 74(19):8953-65.
- Summers DF, Maizel JV Jr. Evidence for large precursor proteins in poliovirus synthesis. *Proc Natl Acad Sci U S A*. 1968 Mar;59(3):966-71.
- Tavaré S. Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences. *Mathematics in the Life Sciences*. 1986. 17: 57–86.
- Taylor MP, Jackson WT. Viruses and arrested autophagosome development. *Autophagy*. 2009. 5(6):870-1.
- Tee KK, Lam TT, Chan YF, Bible JM, Kamarulzaman A, Tong CY, Takebe Y, Pybus OG. Evolutionary genetics of human enterovirus 71: origin, population dynamics, natural selection, and seasonal periodicity of the VP1 gene. *J Virol*. 2010. 84(7):3339-50.
- Thoa le PK, Chiang PS, Khanh TH, Luo ST, Dan TN, Wang YF, Thuong TC, Chung WY, Hung NT, Wang JR, Nhan le NT, Thinh le Q, Su IJ, Dung TD, Lee MS. Genetic and antigenic characterization of enterovirus 71 in Ho Chi Minh City, Vietnam, 2011. *PLoS One*. 2013. 8(7):e69895.
- Thorne JL, Kishino H, Painter IS. Estimating the rate of evolution of the rate of molecular evolution. *Mol Biol Evol*. 1998. 15(12):1647-57.
- Tolskaya EA, Romanova LI, Blinov VM, Viktorova EG, Sinyakov AN, Kolesnikova MS, Agol VI. Studies on the recombination between RNA genomes of poliovirus: the primary structure and nonrandom distribution of crossover regions in the genomes of intertypic poliovirus recombinants. *Virology*. 1987. 161(1):54-61.
- Tsai JD, Kuo HT, Chen SM, Lue KH, Sheu JN. Neurological images and the predictors for neurological sequelae of epidemic herpangina/hand-foot-mouth disease with encephalomyelitis. *Neuropediatrics*. 2014 Apr;45(2):102-8.
- Tu PV, Thao NT, Perera D, Huu TK, Tien NT, Thuong TC, How OM, Cardoso MJ, McMinn PC. Epidemiologic and virologic investigation of hand, foot, and mouth disease, southern Vietnam, 2005. *Emerg Infect Dis*. 2007. 13(11):1733-41.
- Vallet S, Legrand Quillien MC, Dailland T, Podeur G, Gouriou S, Schuffenecker I, Payan C, Marcorelles P. Fatal case of enterovirus 71 infection, France, 2007. *Emerg Infect Dis*. 2009. 15(11):1837-40.
- van de Stolpe A, van der Saag PT. Intercellular adhesion molecule-1. *J Mol Med (Berl)*. 1996 Jan;74(1):13-33. Review.

- van der Sanden S, Koopmans M, Uslu G, van der Avoort H; Dutch Working Group for Clinical Virology. Epidemiology of enterovirus 71 in the Netherlands, 1963 to 2008. *J Clin Microbiol.* 2009. 47(9):2826-33.
- van der Sanden S, van Eek J, Martin DP, van der Avoort H, Vennema H, Koopmans M. Detection of recombination breakpoints in the genomes of human enterovirus 71 strains isolated in the Netherlands in epidemic and non-epidemic years, 1963-2010. *Infect Genet Evol.* 2011. 11(5):886-94.
- Virgen-Slane R, Rozovics JM, Fitzgerald KD, Ngo T, Chou W, van der Heden van Noort GJ, Filippov DV, Gershon PD, Semler BL. An RNA virus hijacks an incognito function of a DNA repair enzyme. *Proc Natl Acad Sci U S A.* 2012. 109(36):14634-9.
- Volle R, Archimbaud C, Couraud PO, Romero IA, Weksler B, Mirand A, Pereira B, Henquell C, Peigue-Lafeuille H, Bailly JL. Differential permissivity of human cerebrovascular endothelial cells to enterovirus infection and specificities of enterovirus 71 in crossing an in vitro model of human blood brain barrier. *J Gen Virol.* 2015 Feb 23. pii: vir.0.000103.
- Volle R, Bailly JL, Mirand A, Pereira B, Marque-Juillet S, Chambon M, Regagnon C, Brebion A, Henquell C, Peigue-Lafeuille H, Archimbaud C. Variations in cerebrospinal fluid viral loads among enterovirus genotypes in patients hospitalized with laboratory-confirmed meningitis due to enterovirus. *J Infect Dis.* 2014. 210(4):576-84.
- Wang B, Xi X, Lei X, Zhang X, Cui S, Wang J, Jin Q, Zhao Z. Enterovirus 71 protease 2Apro targets MAVS to inhibit anti-viral type I interferon responses. *PLoS Pathog.* 2013 Mar;9(3):e1003231.
- Wang SM, Ho TS, Shen CF, Liu CC. Enterovirus 71, one virus and many stories. *Pediatr Neonatol.* 2008. 49(4):113-5.
- Wang SM, Lei HY, Huang KJ, Wu JM, Wang JR, Yu CK, Su IJ, Liu CC. Pathogenesis of enterovirus 71 brainstem encephalitis in pediatric patients: roles of cytokines and cellular immune activation in patients with pulmonary edema. *J Infect Dis.* 2003. 188(4):564-70.
- Wang X, Peng W, Ren J, Hu Z, Xu J, Lou Z, Li X, Yin W, Shen X, Porta C, Walter TS, Evans G, Axford D, Owen R, Rowlands DJ, Wang J, Stuart DI, Fry EE, Rao Z. A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nat Struct Mol Biol.* 2012. 19(4):424-9.
- Wang Y, Feng Z, Yang Y, Self S, Gao Y, Longini IM, Wakefield J, Zhang J, Wang L, Chen X, Yao L, Stanaway JD, Wang Z, Yang W. Hand, foot, and mouth disease in China: patterns of spread and transmissibility. *Epidemiology.* 2011. 22(6):781-92.
- Wang YF, Chou CT, Lei HY, Liu CC, Wang SM, Yan JJ, Su IJ, Wang JR, Yeh TM, Chen SH, Yu CK. A mouse-adapted enterovirus 71 strain causes neurological disease in mice after oral infection. *J Virol.* 2004 Aug;78(15):7916-24.
- Wang YF, Yu CK. Animal models of enterovirus 71 infection: applications and limitations. *J Biomed Sci.* 2014. 17;21:31.
- Ward CD, Flanagan JB. Determination of the poliovirus RNA polymerase error frequency at eight sites in the viral genome. *J Virol.* 1992 Jun;66(6):3784-93.
- Ward T, Pipkin PA, Clarkson NA, Stone DM, Minor PD, Almond JW. Decay-accelerating factor CD55 is identified as the receptor for echovirus 7 using CELICS, a rapid immuno-focal cloning method. *EMBO J.* 1994 Nov 1;13(21):5070-4.
- Wei SH, Huang YP, Liu MC, Tsou TP, Lin HC, Lin TL, Tsai CY, Chao YN, Chang LY, Hsu CM. An outbreak of coxsackievirus A6 hand, foot, and mouth disease associated with onychomadesis in Taiwan, 2010. *BMC Infect Dis.* 2011. 11:346.
- Wei W, Guo H, Li J, Ren S, Wei Z, Bao W, Hu X, Zhao K, Zhang W, Zhou Y, Sun F, Markham R, Yu XF. Circulating HFMD-associated coxsackievirus A16 is genetically and phenotypically distinct from the prototype CV-A16. *PLoS One.* 2014. 9(4):e94746.
- Witsø E, Palacios G, Rønningen KS, Cinek O, Janowitz D, Rewers M, Grinde B, Lipkin WI. Asymptomatic circulation of HEV71 in Norway. *Virus Res.* 2007. 123(1):19-29.

- World Health Organization. 2008. Report on the hand, foot, and mouth disease in Fuyang City, Anhui Province and the prevention and control in China. Beijing, Chinese Center for Disease Control and Prevention and the office of the World Health Organization in China 2008.
- World Health Organization. Enterovirus 71 surveillance. *Wkly. Epidemiol. Rec.* 1979. 7:484-488.
- Xing W, Liao Q, Viboud C, Zhang J, Sun J, Wu JT, Chang Z, Liu F, Fang VJ, Zheng Y, Cowling BJ, Varma JK, Farrar JJ, Leung GM, Yu H. Hand, foot, and mouth disease in China, 2008-12: an epidemiological study. *Lancet Infect Dis.* 2014. 14(4):308-18.
- Yamayoshi S, Iizuka S, Yamashita T, Minagawa H, Mizuta K, Okamoto M, Nishimura H, Sanjoh K, Katsushima N, Itagaki T, Nagai Y, Fujii K, Koike S. 2012. Human SCARB2-dependent infection by coxsackievirus A7, A14, and A16 and enterovirus 71. *J Virol.* 86(10):5686-96.
- Yamayoshi S, Koike S. Identification of a human SCARB2 region that is important for enterovirus 71 binding and infection. *J Virol.* 2011 May;85(10):4937-46.
- Yamayoshi S, Ohka S, Fujii K, Koike S. Functional comparison of SCARB2 and PSGL1 as receptors for enterovirus 71. *J Virol.* 2013 Mar;87(6):3335-47.
- Yamayoshi S, Yamashita Y, Li J, Hanagata N, Minowa T, Takemura T, Koike S. Scavenger receptor B2 is a cellular receptor for enterovirus 71. *Nat Med.* 2009 Jul;15(7):798-801.
- Yang Z, Kumar S, Nei M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics.* 1995. 141(4):1641-50.
- Yang B, Chuang H, Yang KD. Sialylated glycans as receptor and inhibitor of enterovirus 71 infection to DLD-1 intestinal cells. *Virol J.* 2009. ;6:141.
- Yang F, Ren L, Xiong Z, Li J, Xiao Y, Zhao R, He Y, Bu G, Zhou S, Wang J, Qi J. Enterovirus 71 outbreak in the People's Republic of China in 2008. *J Clin Microbiol.* 2009 Jul;47(7):2351-2.
- Yang F, Zhang T, Hu Y, Wang X, Du J, Li Y, Sun S, Sun X, Li Z, Jin Q. Survey of enterovirus infections from hand, foot and mouth disease outbreak in China, 2009. *Virol J.* 2011. 8:508.
- Yang SL, Chou YT, Wu CN, Ho MS. Annexin II binds to capsid protein VP1 of enterovirus 71 and enhances viral infectivity. *J Virol.* 2011. 85(22):11809-20.
- Yang Z, Rannala B. Molecular phylogenetics: principles and practice. *Nat Rev Genet.* 2012. 13(5):303-14.
- Yin XG, Yi HX, Shu J, Wang XJ, Wu XJ, Yu LH. Clinical and epidemiological characteristics of adult hand, foot, and mouth disease in northern Zhejiang, China, May 2008-November 2013. *BMC Infect Dis.* 2014. 14:251.
- Yin-Murphy M. Acute hemorrhagic conjunctivitis. *Prog Med Virol.* 1984. 29:23-44.
- Yip CC, Lau SK, Woo PC, Yuen KY. Human enterovirus 71 epidemics: what's next? *Emerg Health Threats J.* 2013. 6:19780.
- Yip CC, Lau SK, Zhou B, Zhang MX, Tsoi HW, Chan KH, Chen XC, Woo PC, Yuen KY. Emergence of enterovirus 71 "double-recombinant" strains belonging to a novel genotype D originating from southern China: first evidence for combination of intratypic and intertypic recombination events in EV71. *Arch Virol.* 2010. 155(9):1413-24.
- Ylipaasto P, Eskelinen M, Salmela K, Hovi T, Roivainen M. Vitronectin receptors, alpha v integrins, are recognized by several non-RGD-containing echoviruses in a continuous laboratory cell line and also in primary human Langerhans' islets and endothelial cells. *J Gen Virol.* 2010. 91(Pt 1):155-65.
- Yogo Y, Wimmer E. Polyadenylic acid at the 3'-terminus of poliovirus RNA. *Proc Natl Acad Sci U S A.* 1972 Jul;69(7):1877-82.
- Yoke-Fun C, AbuBakar S. Phylogenetic evidence for inter-typic recombination in the emergence of human enterovirus 71 subgenotypes. *BMC Microbiol.* 2006. 6:74.
- Yu H, Chen W, Chang H, Tang R, Zhao J, Gan L, Liu B, Chen J, Wang M. Genetic analysis of the VP1 region of enterovirus 71 reveals the emergence of genotype A in central China in 2008. *Virus Genes.* 2010. 41(1):1-4.
- Zahorsky J. Herpangina. *Am J Dis Child.* 1941. 62(1):169.

- Zhang Y, Tan XJ, Wang HY, Yan DM, Zhu SL, Wang DY, Ji F, Wang XJ, Gao YJ, Chen L, An HQ, Li DX, Wang SW, Xu AQ, Wang ZJ, Xu WB. An outbreak of hand, foot, and mouth disease associated with subgenotype C4 of human enterovirus 71 in Shandong, China. *J Clin Virol*. 2009. 44(4):262-7.
- Zhang Y, Tan XJ, Wang HY, Yan DM, Zhu SL, Wang DY, Ji F, Wang XJ, Gao YJ, Chen L, An HQ, Li DX, Wang SW, Xu AQ, Wang ZJ, Xu WB. An outbreak of hand, foot, and mouth disease associated with subgenotype C4 of human enterovirus 71 in Shandong, China. *J Clin Virol*. 2009. 44(4):262-7.
- Zhang Y, Wang D, Yan D, Zhu S, Liu J, Wang H, Zhao S, Yu D, Nan L, An J, Chen L, An H, Xu A, Xu W. Molecular evidence of persistent epidemic and evolution of subgenotype B1 coxsackievirus A16-associated hand, foot, and mouth disease in China. *J Clin Microbiol*. 2010. 48(2):619-22.
- Zhang Y, Wang D, Yan D, Zhu S, Liu J, Wang H, Zhao S, Yu D, Nan L, An J, Chen L, An H, Xu A, Xu W. Molecular evidence of persistent epidemic and evolution of subgenotype B1 coxsackievirus A16-associated hand, foot, and mouth disease in China. *J Clin Microbiol*. 2010. 48(2):619-22.
- Zhu FC, Meng FY, Li JX, Li XL, Mao QY, Tao H, Zhang YT, Yao X, Chu K, Chen QH, Hu YM, Wu X, Liu P, Zhu LY, Gao F, Jin H, Chen YJ, Dong YY, Liang YC, Shi NM, Ge HM, Liu L, Chen SG, Ai X, Zhang ZY, Ji YG, Luo FJ, Chen XQ, Zhang Y, Zhu LW, Liang ZL, Shen XL. Efficacy, safety, and immunology of an inactivated alum-adjuvant enterovirus 71 vaccine in children in China: a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2013. 381(9882):2024-32.
- Zhu J, Luo Z, Wang J, Xu Z, Chen H, Fan D, Gao N, Ping G, Zhou Z, Zhang Y, An J. Phylogenetic analysis of Enterovirus 71 circulating in Beijing, China from 2007 to 2009. *PLoS One*. 2013. 8(2):e56318.
- Zhu Q, Hao Y, Ma J, Yu S, Wang Y. Surveillance of hand, foot, and mouth disease in mainland China (2008-2009). *Biomed Environ Sci*. 2011. 24(4):349-56.
- Zhu Z, Zhu S, Guo X, Wang J, Wang D, Yan D, Tan X, Tang L, Zhu H, Yang Z, Jiang X, Ji Y, Zhang Y, Xu W. Retrospective seroepidemiology indicated that human enterovirus 71 and coxsackievirus A16 circulated widely in central and southern China before large-scale outbreaks from 2008. *Virology*. 2010. 7:300.
- Zimmerman GA. Two by two: the pairings of P-selectin and P-selectin glycoprotein ligand 1. *Proc Natl Acad Sci U S A*. 2001. 98(18):10023-4.
- Zong W, He Y, Yu S, Yang H, Xian H, Liao Y, Hu G. Molecular phylogeny of Coxsackievirus A16 in Shenzhen, China, from 2005 to 2009. *J Clin Microbiol*. 2011. 49(4):1659-61.

# ANNEXES

# **Manuscrit n°1**

**“Transmission patterns of human enterovirus 71 among European countries, 2003-2013”**

**publié dans le journal « Eurosurveillance » sous forme d’un article de recherche.**

**Accepté le 21 janvier 2015.**

**Auteurs : Chervin Hassel, Audrey Mirand, Alexander Lukashev, Elena Terletskaia-Ladwig, Agnes Farkas, Isabelle Schuffenecker, Sabine Diedrich, Hartwig P. Huemer, Christine Archimbaud, Hélène Peigue-Lafeuille, Cécile Henquell et Jean-Luc Bailly.**



chervin hassel &lt;chervin.hassel@gmail.com&gt;

---

**Fwd: Your manuscript is now accepted -- queued for editing - [EMID:613aa2f4a74b9aba]**

---

Jean-Luc BAILLY &lt;j-luc.bailly@udamail.fr&gt;

21 janvier 2015 17:50

À : chervin hassel &lt;chervin.hassel@gmail.com&gt;

---

**De:** "Eurosurveillance" <em@editorialmanager.com>**À:** "Jean-Luc Bailly" <j-luc.bailly@udamail.fr>**Envoyé:** Mercredi 21 Janvier 2015 17:48:02**Objet:** Your manuscript is now accepted -- queued for editing - [EMID:613aa2f4a74b9aba]

Ref.: Ms. No. eurosurveillance-D-14-00488R1

Transmission patterns of human enterovirus 71 among European countries, 2003-2013  
Eurosurveillance

Dear Dr Bailly,

We have evaluated your revisions and are in principle satisfied with your answers. Where we think something more needs to be done, or more data added, we will point this out to you during the editing stage.

Your manuscript will now enter the queue for editing and publication. You will be contacted when an editor is free.

Please be aware that there will be further questions and clarifications at that stage and you are expected to work with us on the manuscript before publication.

You can now consider your article as accepted. Publication is, however, subject to our questions being answered satisfactorily during editing.

Kind regards,

Dr. Karen Wilson  
Scientific Editor  
Eurosurveillance



## Eurosurveillance

### Transmission patterns of human enterovirus 71 among European countries, 2003-2013

--Manuscript Draft--

<b>Manuscript Number:</b>	eurosurveillance-D-14-00488R1
<b>Full Title:</b>	Transmission patterns of human enterovirus 71 among European countries, 2003-2013
<b>Article Type:</b>	Research
<b>Keywords:</b>	imported viral diseases, laboratory surveillance, outbreaks, enterovirus 71, molecular epidemiology
<b>Corresponding Author:</b>	Jean-Luc Bailly, Ph.D.  FRANCE
<b>Corresponding Author's Institution:</b>	
<b>First Author:</b>	Chervin Hassel
<b>Order of Authors:</b>	Chervin Hassel Audrey Mirand Alexander Lukashev Elena Terletskaia-Ladwig Agnes Farkas Isabelle Schuffenecker Sabine Diedrich Hartwig P Huemer Christine Archimbaud Hélène Peigue-Lafeuille Cécile Henquell Jean-Luc Bailly, Ph.D.
<b>Abstract:</b>	Enterovirus 71 (EV-71) is involved in epidemics of hand, foot, and mouth disease (HFMD) with severe neurological complications in the Asia-Pacific region. The transmission of this virus is poorly understood in other geographic areas. We used large sequence datasets (gene encoding the viral protein 1, VP1) and Bayesian phylogenetic procedure to compare the molecular epidemiology and geographic transportation patterns of EV-71 subgenogroups B4, B5, C1, C2, and C4 in Europe and Asia. The B4, B5, and C4 viruses circulate mainly in Asia with sporadic introductions in Europe. The phylogenies showed evidence of multiple transportation events involving C1 and C2 viruses within Europe since the mid 1990's. The recent occurrence of B5 and C4 infections in Europe and the 2007 Dutch outbreak caused by C2 arose while the transmission of C1 viruses was low. Two waves of C2 infections occurred in 2010 and 2013 in Europe. A transmission chain involving a C4 virus was traced from Japan to Europe and Canada between 2001 and 2006. The patterns of recent virus transportation within and from Europe are involved in the long-term survival of EV-71, which indicates a need for enhanced surveillance of this virus.
<b>Suggested Reviewers:</b>	
<b>Opposed Reviewers:</b>	
<b>Additional Information:</b>	
<b>Question</b>	<b>Response</b>
Word count of the text	3198
Has informed consent has been obtained	Not relevant

from persons whose details are described in your manuscript (or from the persons' guardians) that this information may be published?	
If the answer to the above question is no, explain here why informed consent was not obtained.	
Do you have any conflicts of interest regarding your manuscript?	No
Please copy/paste your covering letter here.	Dear Editors,  Please find attached our manuscript entitled 'Transmission patterns of human enterovirus 71 among European countries'.  We describe a study that assesses and compares the transmission and migration patterns of different subgenogroups of enterovirus 71, in Europe and other geographical regions. The study follows on from previous work because it was based on an expanded virus sample. We show the episodic nature of the enterovirus 71 transmission among European countries, a pattern which contrasts with the epidemic transmission in the Asia Pacific region. The patterns of recent virus transportation within and from Europe may be involved in the long-term survival of EV-71, which indicates a need for enhanced surveillance of this virus.  The manuscript has not been previously published nor is it being considered for publication elsewhere. Parts of the results have been presented at the 18th EUROPIIC (9 – 14 March 2014, Blankenberge, Belgium).  We warrant that all authors have seen and approved the manuscript, and contributed significantly to the work.  We acknowledge the assistance of Jeffrey Watts, who copy-edited the manuscript.  We hope that you will find the manuscript acceptable for consideration in the journal Eurosurveillance.  Yours faithfully, Dr Jean-Luc Bailly
Please copy/paste here the authors' names and affiliations, in the format for publication. The first name of each author should be given in full (i.e. an initial should not be given). Where a collective author is included (e.g. a working group or disease-specific network) and if the persons comprising the group are to be included at the end of the article, please list each person here.	Chervin Hassel <sup>1</sup> , Audrey Mirand <sup>1,2</sup> , Alexander Lukashev <sup>3</sup> , Elena Terletskaia-Ladwig <sup>4</sup> , Agnes Farkas <sup>5</sup> , Isabelle Schuffenecker <sup>6</sup> , Sabine Diedrich <sup>7</sup> , Hartwig P Huemer <sup>8</sup> , Christine Archimbaud <sup>1,2</sup> , Hélène Peigue-Lafeuille <sup>1,2</sup> , Cécile Henquell <sup>1,2</sup> , Jean-Luc Bailly <sup>1,2</sup>  1 Clermont Université, Université d'Auvergne, EPIE, EA 4843, Clermont-Ferrand, France  2 CHU Clermont-Ferrand, Service de Virologie, Centre National de Référence des Entérovirus et Paréchovirus – Laboratoire associé, Clermont-Ferrand, France  3 Chumakov Institute of Poliomyelitis and Viral Encephalites, Moscow, Russia  4 Prof. Gisela Enders & Kollegen MVZ GbR and Institute of Virology, Infectious Diseases and Epidemiology, Stuttgart, Germany  5 Division of Virology, National Center for Epidemiology, Budapest, Hungary  6 Laboratoire de Virologie Est des Hospices Civils de Lyon, Centre National de Référence des Entérovirus et Paréchovirus, Bron, France  7 National Reference Center for Poliomyelitis and Enterovirus, Robert Koch Institute, Berlin, Germany

	8 Austrian Agency for Health and Food Safety, Vienna, Austria 9 Université Claude Bernard, EA4610 VIRPATH, Lyon, France
Authors' contributions: the contribution of each author to the article should be described here. This information will appear at the end of the published article.	Chervin Hassel: performed sequencing and phylogenetic analyses, provided and analysed the data, and provided comments on the manuscript; Audrey Mirand: did isolation and characterisation of virus strains in France, performed sequencing, and analysed the data, and provided comments on the manuscript; Alexander Lukashev: did isolation and characterisation of virus strains, provided sequence data before publication, and provided comments on the manuscript; Elena Terletskaia-Ladwig: did isolation and characterisation of virus strains in Germany, provided clinical data, and provided comments on the manuscript; Agnes Farkas: did isolation and characterisation of virus strains in Hungary, and provided clinical data; Isabelle Schuffenecker: did isolation and characterisation of virus strains in France, provided clinical data; Sabine Diedrich: did isolation and characterisation of virus strains in Germany, provided clinical data; Hartwig P Huemer: did isolation and characterisation of virus strains in Austria, provided clinical data; Christine Archimbaud: did isolation and characterisation of virus strains in France; H�el�ene Peigue-Lafeuille: designed the protocol of investigation, provided comments on the manuscript; C�ecile Henquell: designed the protocol of investigation, analysed the data, provided comments on the manuscript; Jean-Luc Bailly: designed the protocol of investigation, set up the field epidemiology and contacted all investigators, analysed the data, and wrote the first draft.
Please include any acknowledgements here. You may acknowledge anyone who has helped you with any aspect of the report, but it is always the corresponding author's responsibility to obtain permission from anyone being acknowledged. A statement on funding for the work described in the manuscript should be included.	The authors are grateful to the following people involved in the French Network of Enterovirus Surveillance, Dr D. Hecquet and Prof. G. Duverlie (Amiens); Dr A. Paquin, Dr A. Ducancelle, and Prof. F. Lunel-Fabiani (Angers); Dr MC. Legrand and Prof. Christopher Payan (Brest); Dr J. Petitjean and Prof. Astrid Vabret (Caen); Dr Christine Morel-Baccard and Prof. Patrice Morand (Grenoble); Prof. E. Schvoerer (Nancy); Dr M. Coste-Burel (Nantes); Dr A. Bourgoin and Prof. G. Agius (Poitiers); Dr G. Lagathu and Prof. R. Colimon (Rennes); Dr Jean-Michel Mansuy and Prof. Jacques Izopet (Toulouse); Dr St�ephanie Marque-Juillet (Versailles), for providing us with virus samples used in this study. The authors are also grateful to the following people for providing us with virus samples collected in other European countries, Dr A. Marchut and Dr M. Kozmane T�or�ok (National Center for Epidemiology, Budapest, Hungary). The authors are indebted to Prof. Gisela Enders, director of Laboratory Prof. Gisela Enders & Kollegen MVZ (Stuttgart, Germany) for help and constant involvement in our European molecular epidemiology studies. We acknowledge the technical contribution of Gwendoline Jugie, Nathalie Rodde, and Isabelle Simon for helpful assistance with virus genotyping. We thank Jeffrey Watts for help with preparing the English manuscript.
<b>Manuscript Classifications:</b>	Europe; Asia; viral infections; outbreaks; surveillance; epidemiology; molecular methods
<b>Author Comments:</b>	

Ref.: Ms. No. [eurosurveillance-D-14-00488](#)

Transmission patterns of human enterovirus 71 among European countries  
Eurosurveillance

Dear Dr Wilson,

We have addressed all the issues raised by the reviewers.

We provide a point-by-point reply to the comments and a revised version of our manuscript (the changes have been highlighted in yellow in the version with track changes).

In addition, we have rephrased a number of sentences and these changes (requested by neither the reviewers nor the editor) have been highlighted in light grey.

A clean revised version is also provided.

Kind regards,

Dr. Jean-Luc Bailly

1 **Title**

2 **Transmission patterns of human enterovirus 71 among European countries, 2003–2013**

3 **Main text:** 3361 words

4

1 **Abstract (190 words)**

2 Enterovirus 71 (EV-71) is involved in epidemics of hand, foot, and mouth disease (HFMD) with  
3 severe neurological complications in the Asia-Pacific region. The transmission of this virus is  
4 poorly understood in other geographic areas. We used large sequence datasets (gene encoding the  
5 viral protein 1, VP1) and Bayesian phylogenetic procedure to compare the molecular  
6 epidemiology and geographic transportation patterns of EV-71 subgenogroups B4, B5, C1, C2,  
7 and C4 in Europe and Asia. The B4, B5, and C4 viruses circulate mainly in Asia with sporadic  
8 introductions in Europe. The phylogenies showed evidence of multiple transportation events  
9 involving C1 and C2 viruses within Europe since the mid 1990's. The recent occurrence of B5  
10 and C4 infections in Europe and the 2007 Dutch outbreak caused by C2 arose while the  
11 transmission of C1 viruses was low. Two waves of C2 infections occurred in 2010 and 2013 in  
12 Europe. A transmission chain involving a C4 virus was traced from Japan to Europe and Canada  
13 between 2001 and 2006. The patterns of recent virus transportation within and from Europe are  
14 involved in the long-term survival of EV-71, which indicates a need for enhanced surveillance of  
15 this virus.

16

## 1 Introduction

2 The infections caused by enterovirus 71 (EV-71) are associated with hand, foot, and  
3 mouth disease (HFMD) and neurological involvements such as acute meningitis, encephalitis,  
4 and poliomyelitis-like disease [1]. The outbreaks caused by this virus over the last 15 years pose  
5 serious threats to public health in countries of the Asia-Pacific region  
6 ([http://www.wpro.who.int/emerging\\_diseases/HFMD/en/](http://www.wpro.who.int/emerging_diseases/HFMD/en/)) because infections are involved in  
7 cardiopulmonary failure and neurological conditions which, in childhood, can lead to death [2].  
8 During the HFMD epidemics in China between 2010 and 2012, there were 1737 laboratory-  
9 confirmed deaths, of which 93% were associated with EV-71 infections [3]. The occurrence of  
10 fatalities in Asia, notably in China, Malaysia, Vietnam, and Cambodia, does not necessarily  
11 indicate that the EV-71 variants and strains circulating in these countries have a greater virulence  
12 [4]. Rather, it is probably indicative of the total number of infections (HFMD cases and  
13 asymptomatic infections). This infection pattern led health authorities to develop surveillance  
14 systems in most Asian countries. The World Health Organisation provides a guide to clinical  
15 management and outbreak prevention of this emerging disease  
16 ([http://www.wpro.who.int/emerging\\_diseases/documents/HFMDGuidance/en/](http://www.wpro.who.int/emerging_diseases/documents/HFMDGuidance/en/)), for which five  
17 inactivated EV-71 vaccine candidates are being evaluated in clinical trials [5].

18 Unlike in Asia, the EV-71 infections in Europe are not currently subject to particular  
19 epidemiologic surveillance. The virus was involved in outbreaks of neurological diseases in the  
20 1970s in Sweden (1973), Bulgaria (1975; 44 fatalities), Hungary (1978; 47 fatalities), and  
21 France (1979) [6–9]. Recent sero-epidemiological surveys in Germany showed that about 20% of  
22 children aged <5 years, and less than 60% of adults develop neutralizing antibodies against EV-  
23 71, which suggests that mild and mostly undiagnosed EV-71 infections occur in the general  
24 population [10,11]. We and others showed that acute EV-71 infections are regularly investigated

1 in patients admitted to hospital in a number of European countries [12–14]. Severe manifestations  
2 have been occasionally reported in France [15,16], and most documented EV-71 infections over  
3 the last 15 years in European countries resulted in sporadic cases of febrile illness and acute  
4 meningitis [17,18].

5 EV-71 is one serotype among a hundred human enteroviruses (EVs) that belong to the  
6 *Picornaviridae* family. Virus strains can be divided into six genogroups designated A-F [19] but  
7 over the last 50 years only genogroups B and C have been involved in outbreaks. The virus  
8 strains of the two major genogroups are classified into subgenogroups B0 to B5 and C1 to C5 on  
9 the basis of genetic relationships [20]. Subgenogroups B4, B5, and C4 are mainly restricted to  
10 Asian countries while subgenogroups C1 and C2 are chiefly found in Europe. Many aspects of  
11 the epidemiological and evolutionary dynamics of circulating virus strains remain unknown. In  
12 particular, how virus transmission occurs over time, across space, and among genogroups has not  
13 been extensively investigated in geographic regions other than South East Asia [21]. We  
14 investigated the evolutionary dynamics of EV-71 to determine the origin of virus strains sampled  
15 in Europe and their relationships with viruses reported in Asia. To this aim, we used large  
16 sequence datasets of virus isolates sampled over the last 15–30 years, which were mainly  
17 representative of EV-71 subgenogroups circulating in Europe and South East Asia. We estimated  
18 the genetic diversity with the 1D gene encoding the capsid viral protein 1 (VP1; hereafter  
19 designated 1D<sup>VP1</sup>) over time, across geographic regions, and among subgenogroups, something  
20 that has never been done on such a large scale before.

## 22 Materials and methods

### 23 Virus samples, gene amplification, and nucleotide sequencing

1 The enterovirus samples studied consisted of 64 EV-71 strains isolated in Austria, France,  
2 Germany, and Hungary between 2003 and 2013, and 33 strains recovered between 2001 and 2013  
3 in Azerbaijan, Kazakhstan, Kyrgyzstan, and Russia. The EV-71 isolates were collected as part of  
4 routine clinical work-up.

5 The 1D<sup>VP1</sup> sequences were determined in the virus isolates with the genotyping methods  
6 described earlier [14,22]. The complete 1D gene sequence was amplified with either standard  
7 single-round (amplicon length of 1200 bp) or semi-nested (amplicon of 2240 bp) polymerase  
8 chain reaction (PCR) assays. The nucleotide sequences determined in this study were deposited in  
9 the sequence database of the European Molecular Biology Laboratory (EMBL) under the  
10 accession numbers HG934162 – HG934296.

#### 11 **Data collection and compilation of nucleotide sequence datasets**

12 The sequences determined were analysed with sequences obtained from GenBank. The  
13 1D<sup>VP1</sup> sequence datasets were constructed by collating all GenBank entries (as of July 2013 for  
14 1D<sup>VP1</sup> locus) including a sequence of the VP1 capsid protein for any human isolate of EV-71.  
15 Entries reporting nucleotide sequences of <891 nt were discarded and only sequences with fully  
16 specified dates (year) and countries of origin were used. The sequence datasets were constructed  
17 with BioEdit v.7.2.5 software (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) and were  
18 compiled with all sequences determined in our laboratory. The EV-71 1D<sup>VP1</sup> gene sequences  
19 were distributed into five datasets corresponding to subgenogroups B4/B5 (n=217 sequences), C1  
20 (n=280), C2 (n=322), and C4 (n=675). On the basis of earlier phylogenetic data indicating that  
21 subgenogroups B4 and B5 had a common ancestor [21], we analysed jointly their sequences to  
22 increase the sample size. To investigate the EV-71/C4 subgenogroup, an initial dataset of 775  
23 complete sequences was downsized to 675 by removing all sequences but one in genetic clusters  
24 containing multiple sequences with  $\geq 99.5$  % nt identity with one another.

#### 1 **Coalescent estimation of divergence dates and evolutionary rates**

2 Genealogical trees of EV-71 subgenogroups were reconstructed with the BEAST v1.7.5  
3 program [23]. Uncorrelated lognormal prior distributions of substitution rates among lineages  
4 were used for the molecular clock model [24]. The evolutionary history was reconstructed with  
5 the substitution general time reversible (GTR) model. The phylogenetic parameters were co-  
6 estimated in the different analyses by a Markov Chain Monte Carlo (MCMC) process involving  
7  $100 \times 10^6$  to  $200 \times 10^6$  generations to ensure convergence of parameter estimates. The Tracer v.1.5  
8 program (<http://evolve.zoo.ox.ac.uk/software>) was used to check for convergence and mixing  
9 (operator effective sample size >200). The trees estimated with the MCMC procedure were  
10 sampled to obtain a final 20 000 trees. Mean estimates and 95% highest probability density  
11 (HPD) intervals calculated for each operator were compiled by analysing the output files obtained  
12 from the BEAST program with the Tracer v.1.5 program. Maximum clade credibility (MCC)  
13 trees were calculated with the TreeAnnotator v.1.7.5 program  
14 (<http://evolve.zoo.ox.ac.uk/software>) and topological support was assessed by estimating the  
15 values of the posterior probability (pp) density of each node.

#### 16 **Reconstruction of the demographic history and geographic spread of EV subgenogroups**

17 The Bayesian skyline plot model (BSP) does not assume a pre-defined model of  
18 demography [25]. It infers a demographic parameter representing 'virus population size', also  
19 referred to as relative genetic diversity, using the coalescent theory and temporal information of  
20 the molecular data (virus collection date). Genetic diversity is the product of the effective size of  
21 the virus population ( $N_e$ ) and the generation time ( $t$ ), and reflects the effective number of  
22 infections averaged over time under the assumption of a neutral evolutionary process. In our  
23 study, this model was used to investigate possible variations in virus population size over time for  
24 subgenogroups B4, B5, C1, C2, and C4 to describe past epidemiological events on a continuous

1 time scale using large sequence datasets for the 1D<sup>VP1</sup> gene. EV-71 infections are characterised  
2 by epidemic transmission in a number of countries, a high rate of asymptomatic or pauci-  
3 symptomatic infections, and short-lived acute infections in susceptible individuals. Accordingly,  
4 the skyline plot model appears epidemiologically better suited than the other demographic models  
5 (constant population size, exponential growth, logistic growth, and expansion growth) for  
6 modelling the spread of EV-71 populations and thus for reconstructing the diffusion of HFMD  
7 over time and across geographic regions. The BSP analyses were done with a lognormal  
8 distribution (piecewise-constant population size model) to account for variation in substitution  
9 rates among the phylogenies. The countries where the virus strains were collected were used as  
10 discrete character states in the phylogeographic analysis to estimate changes of geographic  
11 locations on the phylogenetic trees [26]. The geographic locations of the ancestral nodes were co-  
12 estimated with the phylogeny by using a discrete phylogeographic diffusion model (symmetrical  
13 substitution model). Bayesian stochastic search variable selection was also used for identifying  
14 the statistically significant transition rates between locations. TreeAnnotator software was used to  
15 produce the MCC tree, with the branches coloured by traits.

16

## 17 **Results**

### 18 **Phylogenies of five EV-71 subgenogroups estimated with the 1D<sup>VP1</sup> datasets**

19 We used the large number of 1D<sup>VP1</sup> sequences determined in circulating strains to obtain a  
20 comprehensive comparison of virus transmission among EV-71 subgenogroups. The country of  
21 virus isolation was checked for each sequence to determine the distribution of four main  
22 geographic locations (Asia, Europe, North America, and Russia) as observed in the genealogies.  
23 The phylogenies displayed a 'ladder-like' shape characterised by a long trunk with short side  
24 branches, evidence of rapid extinction of lineages over time (Figure 1). These topologies arose

1 from a combined effect of temporal sampling in various countries and rapid coalescence. A  
2 number of long side branches were also inferred in the C1, C2, and C4 trees, which indicated  
3 transmission of lineages over time and the persistence of these lineages alongside more recent  
4 ones. All but one of the subgenogroups examined arose in the 1990s (Figures 1A, 1C, and 1D).  
5 The time to most recent common ancestor (tMRCA) of subgenogroup EV-71/C1 was estimated in  
6 1983.5 (95% HPD, 1982.2–1984.4) (Figure 1B). Most B4, B5 and C4 lineages had geographic  
7 sampling in East and South East Asia while subgenogroups C1 and C2 showed mixing of  
8 geographic origins between Asia, Europe, and Russia over the sampling periods. The C4  
9 phylogeny also included a few lineages in Europe and Russia, which suggests sporadic virus  
10 transmission from Asia.

### 11 **Genealogy-based population dynamics of EV-71 subgenogroups**

12 The demographic histories showed variations between the five subgenogroups. The pattern  
13 estimated by our coalescent-based reconstruction for subgenogroup C4 clearly indicated an  
14 annual series of increases in genetic diversity interspersed with genetic bottlenecks (data not  
15 shown). This pattern was caused by the large HFMD epidemics recently observed in China [3]  
16 and confirmed the usefulness of the BSP model for investigating EV-71 transmission. We provide  
17 the detailed data obtained for the most frequent subgenogroups (C1 and C2) in Europe (Figure 2).  
18 The pattern estimated for subgenogroup C1 (Figure 2A) depicted successive variations in  
19 population size. This demographic pattern was consistent with a few outbreaks over time and  
20 across different geographic regions as documented by epidemiological data [20, 27]. This pattern  
21 may also indicate uninterrupted global circulation of the virus. The pattern reconstructed for EV-  
22 71/C2 disclosed two distinct rises in population size (Figure 2B). The subgenogroup experienced  
23 a first sharp increase in 1998 that coincided with the occurrence of a large HFMD epidemic in  
24 Taiwan [28]. The population size then slowly decreased over several years. The second

1 exponential rise in virus population size (in 2007) was of a high magnitude. This major increase  
2 in virus genetic diversity coincided with the occurrence of an outbreak in the Netherlands [29]  
3 and a concomitant rise in virus circulation in several European countries [14]. The transmission  
4 of the virus remained at a sustained high level after the 2007 epidemic. A combination of the C1  
5 and C2 patterns showed that the decrease in the EV-71/C1 population size in 1996 was  
6 concurrent with the initial occurrence of EV-71/C2 (Figure 2C). The 1998 increase in population  
7 size (Taiwan epidemic, subgenogroup C2, peak 1) coincided with a lower genetic diversity of  
8 EV-71/C1. The C1 pattern showed three yearly ascending rises (peaks 2, 3, and 4) during a long  
9 period of low genetic diversity of the C2 virus. The population size of both subgenogroups  
10 remained low over the years 2004–2006, a period during which the EV-71/C4 subgenogroup  
11 occurred in several European countries and Canada (see below). The marked increase in C2  
12 genetic diversity in 2007 (peak 5) occurred in Europe during a period of low C1 diversity.

#### 13 **Transportation of EV-71/C1 between distinct countries**

14 EV-71 transmission between countries was investigated with the MCC trees of the above  
15 analyses. A phylogenetic pattern indicative of a probable virus transportation event between two  
16 nodes was defined as follows: (1) the nodes exhibited  $pp > 0.9$ , (2) the inferred location  
17 probabilities were  $> 0.7$  at both nodes, and (3) the difference between the TMRCA values or the  
18 95% HPD interval values of nodes were in a range of one year. With this conservative approach,  
19 the MCC tree topology inferred for subgenogroup C1 indicated a total of 13 consistent virus  
20 transmission events (Figure 3A). Eight phylogenetic patterns (numbered 03–05, 07–09, 11 and  
21 12) indicated virus transmission events between various European countries. Great Britain was  
22 involved as a source country with a probability of 0.98 in three virus transportation events (03,  
23 08, and 12) to France (probabilities of 0.95) and one transmission to Spain (0.99). France was the  
24 source of two transmission events, one event to each country Germany and Great Britain. The

1 EV-71/C1 tree also showed that virus transmissions occurred from Austria (location probability,  
2 0.83) to Germany and from the Netherlands (0.84) to Finland. Three other phylogenetic patterns  
3 (numbered 06, 10, and 13) were indicative of long distance transmission events from a European  
4 country to either Australia or Japan ( $n=2$  events). In the transmission event number 13, the  
5 probability of Great Britain being the source country was estimated to be 0.56 as against 0.28 for  
6 France. Two consistent phylogenetic patterns showed virus transmissions from Malaysia  
7 (location probability  $\geq 0.94$ ): one transportation event spread the virus to Singapore and the other  
8 to Thailand. A schematic representation of all virus transportation events described above is  
9 shown in Figure 4.

#### 10 **Transportation of EV-71/C2 between distinct countries**

11 In the MCC tree inferred for subgenogroup C2 (Figure 3B), 12 phylogenetic patterns  
12 indicated virus transmission events between distinct countries within Western Europe, Georgia,  
13 and Russia. The most frequent source in the inferred virus transportation events was the  
14 Netherlands ( $n=5$  events; location probability range, 0.87–1) and France ( $n=5$ ; 0.74–0.98)  
15 followed by Finland and Germany. Three virus transmissions were estimated between distant  
16 countries: from Finland to Georgia and from France to Russia. An additional transmission event  
17 from Japan (location probability of 0.85) to Spain was also inferred in the genealogy before 1995  
18 (not shown in Figure 3B). All the virus transportation events are shown in Figure 4.

#### 19 **Transportation of EV-71/C4 between distinct countries**

20 We analysed separately a sequence subset corresponding to the lineage highlighted in  
21 Figure 1D. The distributions of geographic location probabilities were indicated for the most  
22 relevant nodes in the MCC tree inferred for this sequence subset and were used to investigate  
23 virus transportations between countries (Figure 5). The phylogeny pattern indicated that between  
24 2001 and 2002 the virus strain was transported from China to Japan, and from there to Europe in

1 early 2003. The virus was disseminated in different European countries between 2003 and 2004.  
2 It was sampled in Croatia in 2005 and the phylogenetic analysis estimated a probability of 0.92  
3 that the source country was Germany. The same virus strain also moved from Europe to Canada  
4 (sampling year 2006). The phylogenetic analysis suggested that the virus was transported from  
5 Germany with a probability of 0.72.

## 6 **Discussion**

7 The recent reporting of sporadic circulation of EV-71 strains of the 'Asian  
8 subgenogroups' B5 [30] and C4 [31–33] in European countries makes it important to determine  
9 the evolutionary origins and spatiotemporal spread of EV-71 infections in geographic areas other  
10 than the Asia-Pacific region. There is a large body of literature on the molecular epidemiology of  
11 EV-71 [34] but, to our knowledge, the spread of this virus has not yet been analysed explicitly  
12 with recently developed phylogeographic models [26]. In this report, we therefore applied  
13 demographic and phylogeographic analyses to compare the population structure and the  
14 dissemination patterns of EV-71 subgenogroups in Europe and Asia. We found that the EV-71  
15 subgenogroups exhibit distinct phylodynamic patterns and that these patterns differed from those  
16 of past epidemics in frequency and geographic location. The transmission history of  
17 subgenogroup C2 was characterised by fewer increases in genetic diversity than in  
18 subgenogroups C1 and C4. For instance, the genetic diversity pattern of subgenogroup C2 shows  
19 clearly the two outbreaks (Taiwan, 1998 and the Netherlands, 2007) documented in the literature  
20 [28, 29]. The variations in genetic diversity over the study period also reflect gaps in virus  
21 sampling over time and surveillance differences between countries. In Europe, EV-71 sequences  
22 are determined solely in patients admitted to hospitals and in a few countries only, whereas in a  
23 number of Asian countries real-time epidemiological data are obtained from sentinel surveillance  
24 of HFMD cases. In this study, we used only full-length EV-71 1D<sup>VP1</sup> sequences because

1 phylogenetic trees based on short-length sequences did not provide a high level of statistical  
2 confidence. Partial nt sequences were notably determined for virus genotyping and ranged over  
3 different parts of the 1D<sup>VP1</sup> gene. Accordingly, we had to discard a certain amount of sequence  
4 data, some from Europe, which might have been helpful in phylogenetic analysis of the virus  
5 transmission patterns.

6 We show how transmission of EV-71 strains (subgenogroups C1, C2, and C4) occurs in  
7 Europe as discrete and temporally defined virus introductions occasionally followed by limited  
8 local spread. The only exception to this pattern was the epidemic expansion in the Netherlands in  
9 2007. Interestingly, we found limited phylogenetic evidence that the Dutch outbreak was the  
10 source of a large spread to other European countries. The phylogenetic patterns also show that the  
11 European region may experience multiple virus introduction events within the same year. This  
12 dissemination mode was observed within particular countries (e.g. France, Germany) for both  
13 subgenogroups C1 and C2 but was more clearly seen during three waves of sporadic and  
14 epidemic infections in 2007, 2010, and 2013 (subgenogroup C2). The epidemiological and  
15 biological factors involved are unknown but the occurrence of these infection waves is consistent  
16 with the hypothesis that the sustained circulation of an EV-71 strain throughout Europe depends  
17 on the proportion of susceptible hosts in different countries. The data may alternatively indicate  
18 that the immunity elicited by the C1 and C2 infections is cross-protective, as suggested by earlier  
19 studies that identified common epitopes and suggested that human immune sera among virus  
20 strains of different EV-71 subgenogroups have cross-neutralization properties [35]. In this  
21 respect, we also provide phylogenetic evidence that the spread of the C2 virus across European  
22 countries in 2007 happened at a time when C1 infections had been transmitted at low rates for at  
23 least three years.



1 It is particularly noteworthy that the transmission of EV-71 strains C1 and C2 in Europe is  
2 mainly dependent on the frequency of virus transportation events between neighbouring  
3 countries. This has been suggested for EV-71 in Asia [36] and for another EV, coxsackievirus B5  
4 [37]. The present study also indicates that the long-term survival of EV-71 (C1 and C2) depends  
5 on continued virus transmission across larger geographic regions, notably Russia and Asia. We  
6 also reconstructed a consistent transmission chain caused by a C4 virus strain throughout Europe  
7 (Austria, Croatia, France, Germany, and Hungary), which suggests that the virus persisted  
8 between 2003 and 2005. The chain was brought about by a virus from Japan which eventually  
9 reached Canada in 2005. Transmission of the C1 and C2 infections was low during the whole  
10 period in Europe. Similarly, the recent C4 infections in France (2012), and Russia (2013) arose  
11 during a transmission trough of both C1 and C2 infections. Thus, transcontinental transmission  
12 events should be considered in the global epidemiology of the virus: they provide the  
13 epidemiological bases for explaining the long-term survival of lineages despite abrupt extinctions  
14 that drastically reduce virus diversity after the occurrence of outbreaks in particular locations.

15 Earlier hypotheses proposed that yearly epidemics of EV-71 arose from transmission  
16 between neighbouring countries in South East and South Pacific Asia. We suggest that virus  
17 strains circulate outside these regions and their spread is sustained by complex migration  
18 dynamics involving Europe and Russia, and possibly other geographic regions. This argument is  
19 lent weight by a recent study reporting the occurrence of EV-71 C4 in Denmark in 2012 and 2013  
20 [38]. Our Bayesian analysis of the partial 1D<sup>VP1</sup> sequences showed that the C4 viruses isolated in  
21 Denmark were distributed in two distinct phylogenetic clusters, which suggests independent virus  
22 sources (data not shown). Thus, the frequency of epidemics in Asian countries, the involvement  
23 of EV-71 in neurological conditions, the recent isolation of C4 variant strains in Europe, and the  
24 transportation patterns of the virus between distant and neighbouring countries underline a need

1 for enhanced surveillance of EV-71 infections in Europe using common enterovirus genotyping  
2 methods.  
3

## 1 References

- 2 **1.** Ooi MH, Wong SC, Lewthwaite P, Cardoso MJ, Solomon T. Clinical features, diagnosis  
3 and management of human enterovirus 71 infection. *Lancet Neurol.* 2010;9(11):1097–1105.  
4
- 5 **2.** Ho M, Chen ER, Hsu KH, Twu SJ, Chen KT, Tsai SF, et al. An epidemic of enterovirus  
6 71 infection in Taiwan. *N Engl J Med.* 1999;341(13):929–35.  
7
- 8 **3.** Xing W, Liao Q, Viboud C, Zhang J, Sun J, Wu JT, et al. Hand, foot, and mouth disease  
9 in China, 2008-12: an epidemiological study. *Lancet Infect Dis.* 2014;14(4):308–18.  
10
- 11 **4.** Chang LY, King CC, Hsu KH, Ning HC, Tsao KC, Li CC, et al. Risk factors of  
12 enterovirus 71 infection and associated hand, foot, and mouth disease/herpangina in children  
13 during an epidemic in Taiwan. *Pediatrics.* 2002;109(6):e88.  
14
- 15 **5.** Lee MS, Tseng FC, Wang JR, Chi CY, Chong P, Su IJ. Challenges to licensure of  
16 enterovirus 71 vaccines. *PLoS Negl Trop Dis.* 2012;6(8):e1737.  
17
- 18 **6.** Sohier R. Enterovirus 71 surveillance: France. *World Health Organ Wkly Epidemiol Rec.*  
19 1979;54(28):219.  
20
- 21 **7.** Chumakov M, Voroshilova M, Shindarov L, Lavrova I, Gracheva L, Koroleva G, et al.  
22 Enterovirus 71 isolated from cases of epidemic poliomyelitis-like disease in Bulgaria. *Arch Virol.*  
23 1979;60(3-4):329–40.  
24
- 25 **8.** Shindarov LM, Chumakov MP, Voroshilova MK, Bojinov S, Vasilenko SM, Iordanov I,  
26 et al. Epidemiological, clinical, and pathomorphological characteristics of epidemic  
27 poliomyelitis-like disease caused by enterovirus 71. *J Hyg Epidemiol Microbiol Immunol.*  
28 1979;23(3):284–95.  
29
- 30 **9.** Blomberg J, Lycke E, Ahlfors K, Johnsson T, Wolontis S, von Zeipel G. New enterovirus  
31 type associated with epidemic of aseptic meningitis and/or hand, foot, and mouth disease. *Lancet*  
32 1974;2(7872):112.  
33
- 34 **10.** Diedrich S, Weinbrecht A, Schreier E. Seroprevalence and molecular epidemiology of  
35 enterovirus 71 in Germany. *Arch Virol.* 2009;154(7):1139–42.  
36
- 37 **11.** Rabenau HF, Richter M, Doerr HW. Hand, foot and mouth disease: seroprevalence of  
38 Coxsackie A16 and Enterovirus 71 in Germany. *Med Microbiol Immunol.* 2010;199(1):45–51.  
39
- 40 **12.** Roth B, Enders M, Arents A, Pfitzner A, Terletskaia-Ladwig E. Epidemiologic aspects  
41 and laboratory features of enterovirus infections in Western Germany,  
42 2000-2005. *J Med Virol.* 2007;79(7):956–62.  
43
- 44 **13.** van der Sanden S, Koopmans M, Uslu G, van der Avoort H; Dutch Working Group for  
45 Clinical Virology. Epidemiology of enterovirus 71 in the Netherlands, 1963 to 2008. *J Clin*  
46 *Microbiol.* 2009;47(9):2826–33.

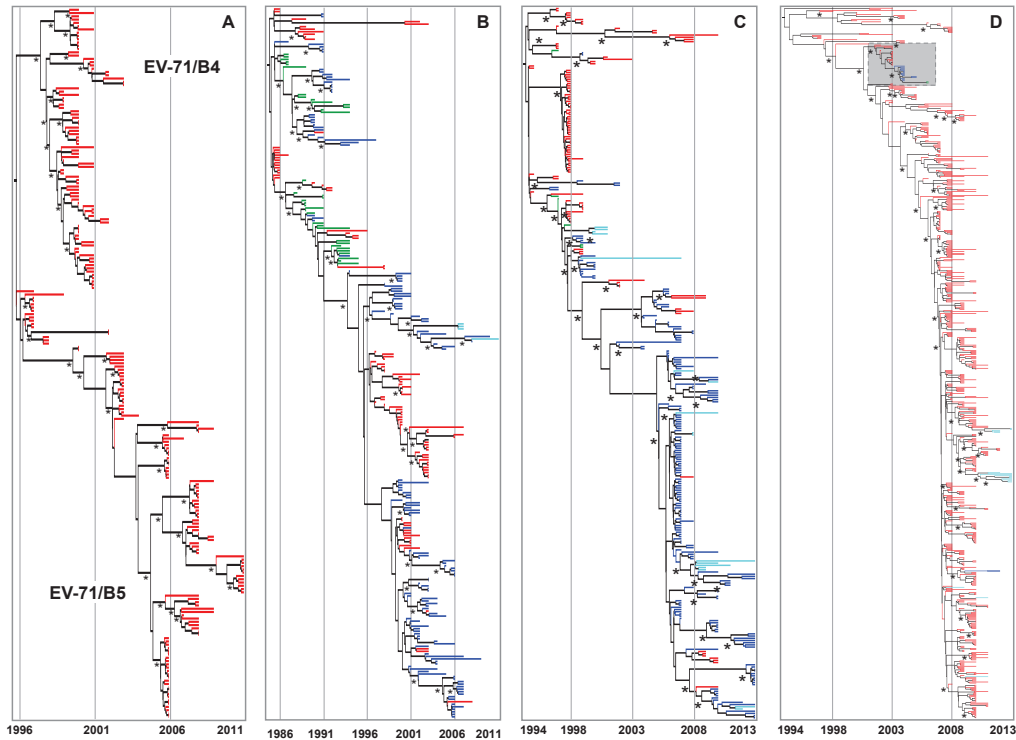
## 1

- 2 **14.** Mirand A, Schuffenecker I, Henquell C, Billaud G, Jugie G, Falcon D, et al. Phylogenetic  
3 evidence for a recent spread of two populations of human enterovirus 71 in European countries. *J*  
4 *Gen Virol.* 2010;91(9):2263–77.  
5
- 6 **15.** Vallet S, Legrand Quillien MC, Dailland T, Podeur G, Gouriou S, Schuffenecker I, et al.  
7 Fatal case of enterovirus 71 infection, France, 2007. *Emerg Infect Dis.* 2009;15(11):1837–40.  
8
- 9 **16.** Kassab S, Saghi T, Boyer A, Lafon ME, Gruson D, Lina B, et al. Fatal case of enterovirus  
10 71 infection and rituximab therapy, France, 2012. *Emerg Infect Dis.* 2013;19(8):1345–1347.  
11
- 12 **17.** Schuffenecker I, Mirand A, Antona D, Henquell C, Chomel JJ, Archimbaud C, et al.  
13 Epidemiology of human enterovirus 71 infections in France, 2000-2009. *J Clin Virol.*  
14 2011;50(1):50–6.  
15
- 16 **18.** Honkanen H, Oikarinen S, Pakkanen O, Ruokoranta T, Pulkki MM, Laitinen OH, et al.  
17 Human enterovirus 71 strains in the background population and in hospital patients in Finland. *J*  
18 *Clin Virol.* 2013;56(4):348–53.  
19
- 20 **19.** Bessaud M, Razafindratsimandresy R, Nougairède A, Joffret ML, Deshpande JM, Dubot-  
21 Pérès A, et al. 2014. Molecular comparison and evolutionary analyses of VP1 nucleotide  
22 sequences of new African Human enterovirus 71 isolates reveal a wide genetic diversity. *PLOS*  
23 *One.* 2014;9(3):e90624.  
24
- 25 **20.** McMinn P. Recent advances in the molecular epidemiology and control of human  
26 enterovirus 71 infection. *Cur Opin Virol.* 2012;2(2):199–205.  
27
- 28 **21.** Tee KK, Lam TT, Chan YF, Bible JM, Kamarulzaman A, Tong CY, et al. Evolutionary  
29 genetics of human enterovirus 71: origin, population dynamics, natural selection, and seasonal  
30 periodicity of the VP1 gene. *J Virol.* 2010;84(7):3339–50.  
31
- 32 **22.** Mirand A, Henquell C, Archimbaud C, Ughetto S, Antona D, Bailly JL, et al. Outbreak of  
33 hand, foot and mouth disease/herpangina associated with coxsackievirus A6 and A10 infections  
34 in 2010, France: a large citywide, prospective observational study. *Clin Microbiol Infect.*  
35 2012;18(5):E110–8.  
36
- 37 **23.** Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti  
38 and the BEAST 1.7. *Mol Biol Evol.* 2012;29(8):1969–73.  
39
- 40 **24.** Drummond AJ, Ho SY, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with  
41 confidence. *PLoS Biol.* 2006;4(5):e88.  
42
- 43 **25.** Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past  
44 population dynamics from molecular sequences. *Mol Biol Evol.* 2005;22(5):1185–92.  
45
- 46 **26.** Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its  
47 roots. *PLoS Comput Biol.* 2009;5(9):e1000520.

- 1  
2 **27.** Yip CC, Lau SK, Woo PC, Yuen KY. Human enterovirus 71 epidemics: what's next?  
3 *Emerg Health Threats J.* 2013;6:19780.  
4  
5 **28.** Wang JR, Tuan YC, Tsai HP, Yan JJ, Liu CC, Su IJ. Change of major genotype of  
6 enterovirus 71 in outbreaks of hand-foot-and-mouth disease in Taiwan between 1998 and 2000. *J*  
7 *Clin Microbiol.* 2002;40(1):10–5.  
8  
9 **29.** van der Sanden S, van der Avoort H, Lemey P, Uslu G, Koopmans M. Evolutionary  
10 trajectory of the VP1 gene of human enterovirus 71 genogroup B and C viruses. *J Gen Virol.*  
11 2010;91(8):1949–58.  
12  
13 **30.** Badran SA, Midgley S, Andersen P, Böttiger B. Clinical and virological features of  
14 enterovirus 71 infections in Denmark, 2005 to 2008. *Scand J Infect Dis.* 2011;43(8):642–8.  
15  
16 **31.** Ortner B, Huang CW, Schmid D, Mutz I, Wewalka G, Allenberger F, et al. Epidemiology  
17 of enterovirus types causing neurological disease in Austria 1999-2007: detection of clusters of  
18 echovirus 30 and enterovirus 71 and analysis of prevalent genotypes. *J Med Virol.*  
19 2009;81(2):317–24.  
20  
21 **32.** Kapusinszky B, Szomor KN, Farkas A, Takács M, Berencsi G. Detection of non-polio  
22 enteroviruses in Hungary 2000-2008 and molecular epidemiology of enterovirus 71,  
23 coxsackievirus A16, and echovirus 30. *Virus Genes.* 2010;40(2):163–73.  
24  
25 **33.** Schuffenecker I, Henquell C, Mirand A, Coste-Burel M, Marque-Juillet S, Desbois D, et  
26 al. New introductions of enterovirus 71 subgenogroup C4, France, 2012. *New Introductions of*  
27 *Enterovirus 71 Subgenogroup C4 Strains, France, 2012. Emerg Infect Dis.* 2014;20(8):13436.  
28  
29 **34.** Solomon T, Lewthwaite P, Perera D, Cardoso MJ, McMinn P, Ooi MH. Virology,  
30 epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect Dis.* 2010;10(11):778–  
31 90.  
32  
33 **35.** Liu CC, Chou AH, Lien SP, Lin HY, Liu SJ, Chang JY, et al. Identification and  
34 characterization of a cross-neutralization epitope of enterovirus 71. *Vaccine.* 2011;29(26):4362–  
35 72.  
36  
37 **36.** Mizuta K, Abiko C, Murata T, Matsuzaki Y, Itagaki T, Sanjoh K, et al. Frequent  
38 importation of enterovirus 71 from surrounding countries into the local community of Yamagata,  
39 Japan, between 1998 and 2003. *J Clin Microbiol.* 2005;43(12):6171–5.  
40  
41 **37.** Henquell C, Mirand A, Richter J, Schuffenecker I, Böttiger B, Diedrich S, et al.  
42 Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two  
43 genogroups and reveal evolutionary factors of molecular adaptation and transmission. *J Virol.*  
44 2013;87(22):12249–59.  
45  
46 **38.** Fischer T, Nielsen A, Sydenham T, Andersen P, Andersen B, Midgley S. Emergence of  
47 enterovirus 71 C4a in Denmark, 2009 to 2013. *Euro Surveill.* 2014;19(38):pii=20911.

Figure

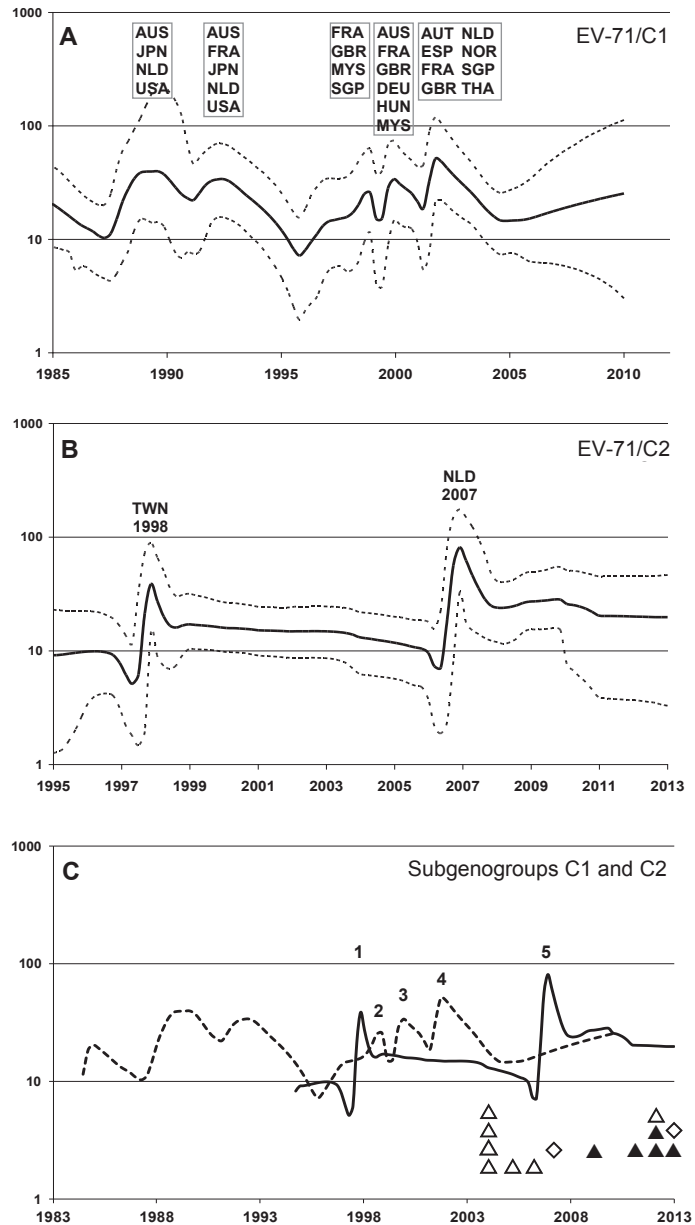
Figure 1



**Figure 1** Comparison of phylogenies of five EV-71 subgenogroups showing geographic origins of sample sequences

The chronogram trees were inferred with the 1D<sup>VP1</sup> gene sequence datasets for subgenogroups B4/B5 (A), C1 (B), C2 (C), and C4 (D). The tree topologies show the genetic relationships between taxa sampled over the periods indicated on the x axis (calendar years). The phylogenetic relationships were inferred with a Bayesian method using a relaxed molecular-clock model. For clarity, the sequence names are not included in the tree. Asterisks indicate key nodes with posterior probability (pp) density values > 0.90. Each tip branch represents a sampled virus sequence. The branches in the genealogies of each subgenogroup were colored to investigate relationships between phylogenetic clustering and the geographic origins of taxa. The geographic locations where the virus strains were sampled are indicated by different colors: blue, Europe; light blue, Russia and associated States; green, North America; and red, Asia Pacific region.

Figure 2



**Figure 2** Plots of genetic diversity over time (calendar years) reconstructed for the EV-71 subgenogroups C1 (A), C2 (B), and C1/C2 (C)

Global genetic diversity over time was estimated with gene  $1D^{VP1}$  for each subgenogroup. The x axis indicates time (calendar years). The y axis indicates genetic diversity ( $\log_{10}N_e t$ ;  $N_e$ , effective size of virus population;  $t$ , generation time) estimated with the Bayesian skyline plot method. The mean value of the effective number of infections is shown with a continuous line. The upper and lower 95% highest probability density intervals are shown with dotted lines. The plots estimated for the EV-71 subgenogroups C1 and C2 were superimposed on the same time scale in Figure 2C and are represented with dotted and solid lines, respectively. The epidemic peaks are numbered. The introductions of EV-71/C4 strains in Europe and Russia are indicated with open and full triangles, respectively. The introduction events of the EV-71 subgenogroup B5 in Denmark and France (data not shown) are shown with diamonds. The countries where the virus strains were detected frequently are indicated above the peaks of genetic diversity. Abbreviations used: AUS, Australia; AUT, Austria; DEU, Germany; ESP, Spain; FRA, France; GBR, Great Britain; HUN, Hungary; JPN, Japan; MYS, Malaysia; NLD, Netherlands; NOR, Norway; SGP, Singapore; THA, Thailand; USA, United states of America.

Figure 3A

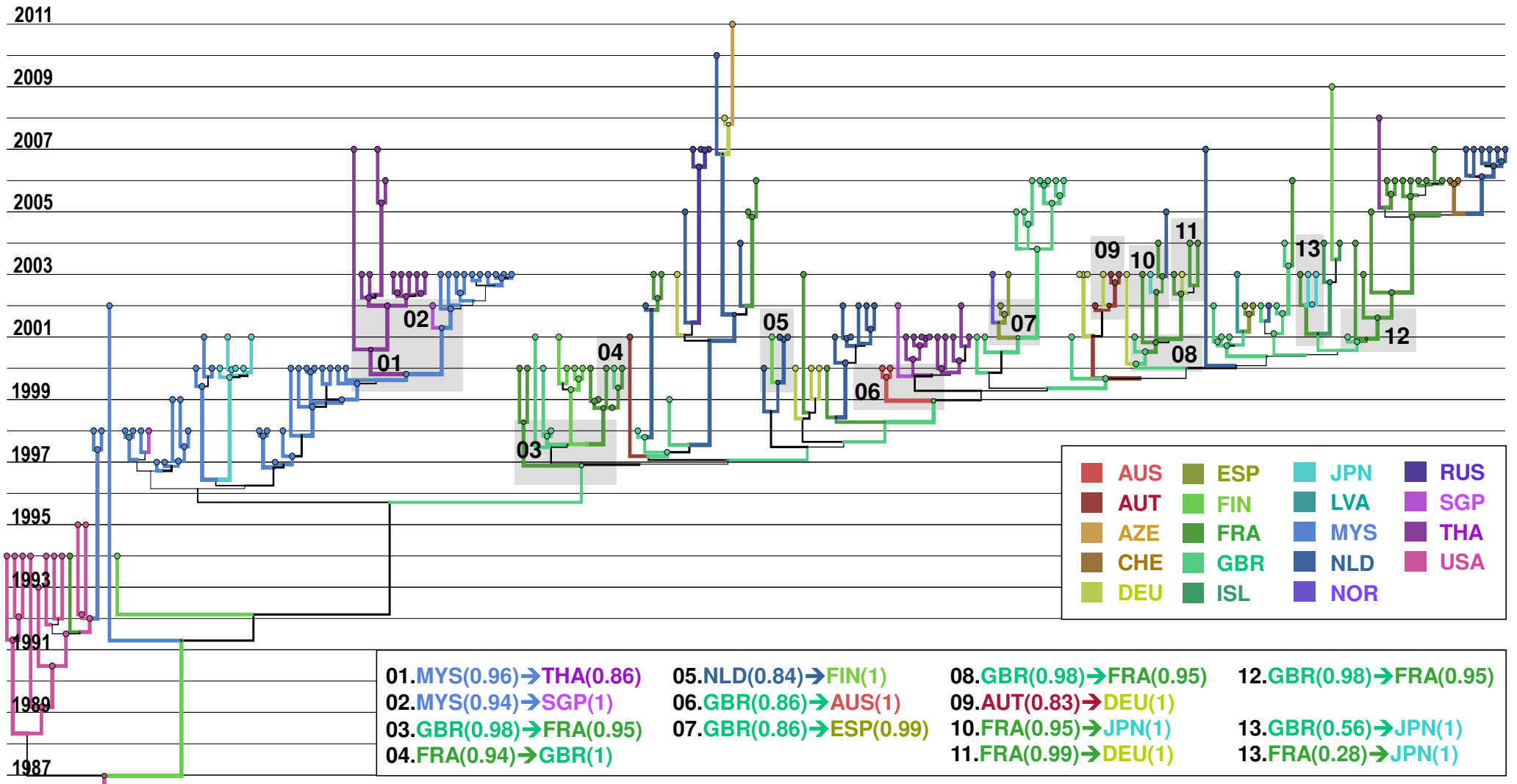
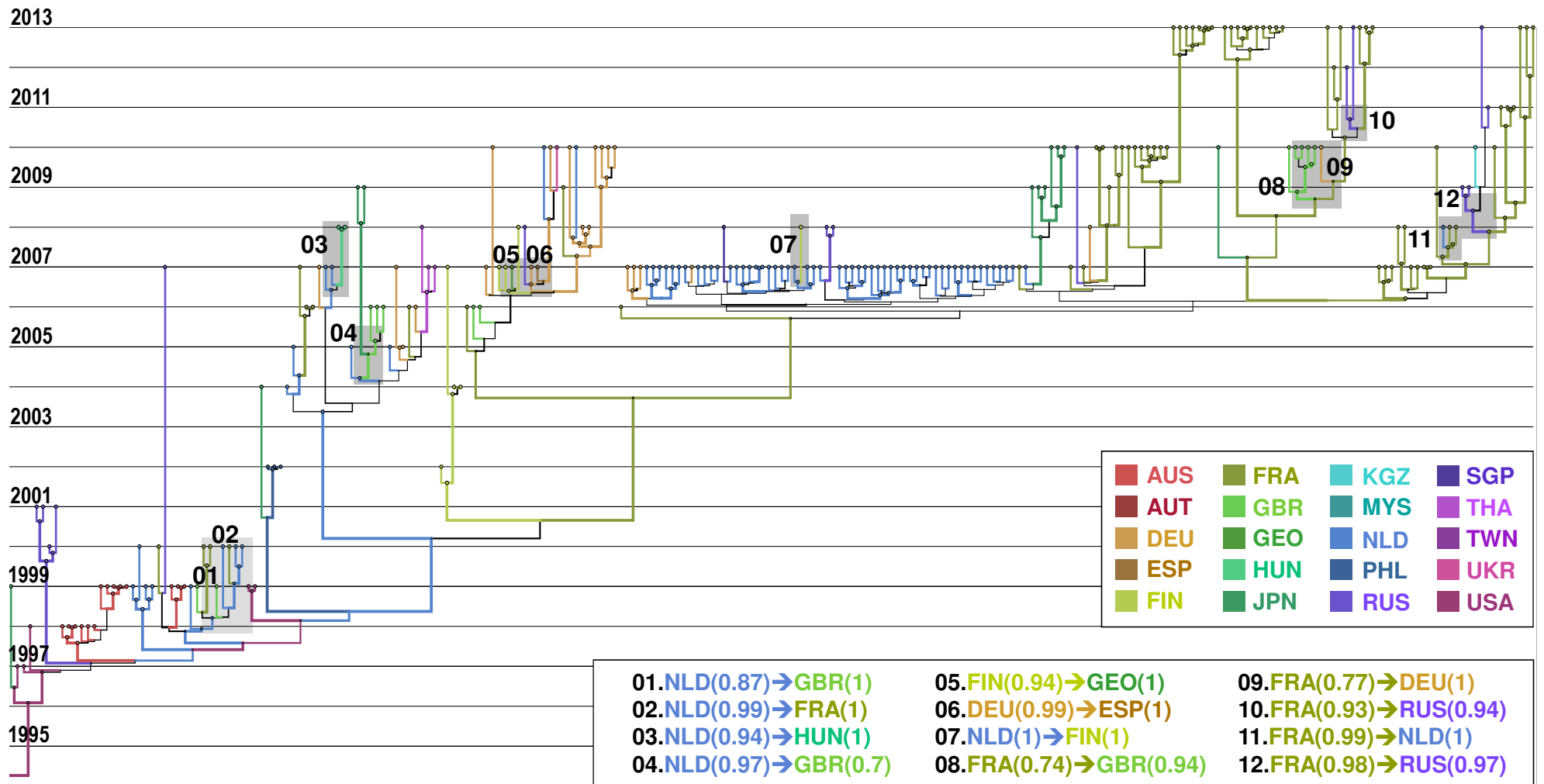


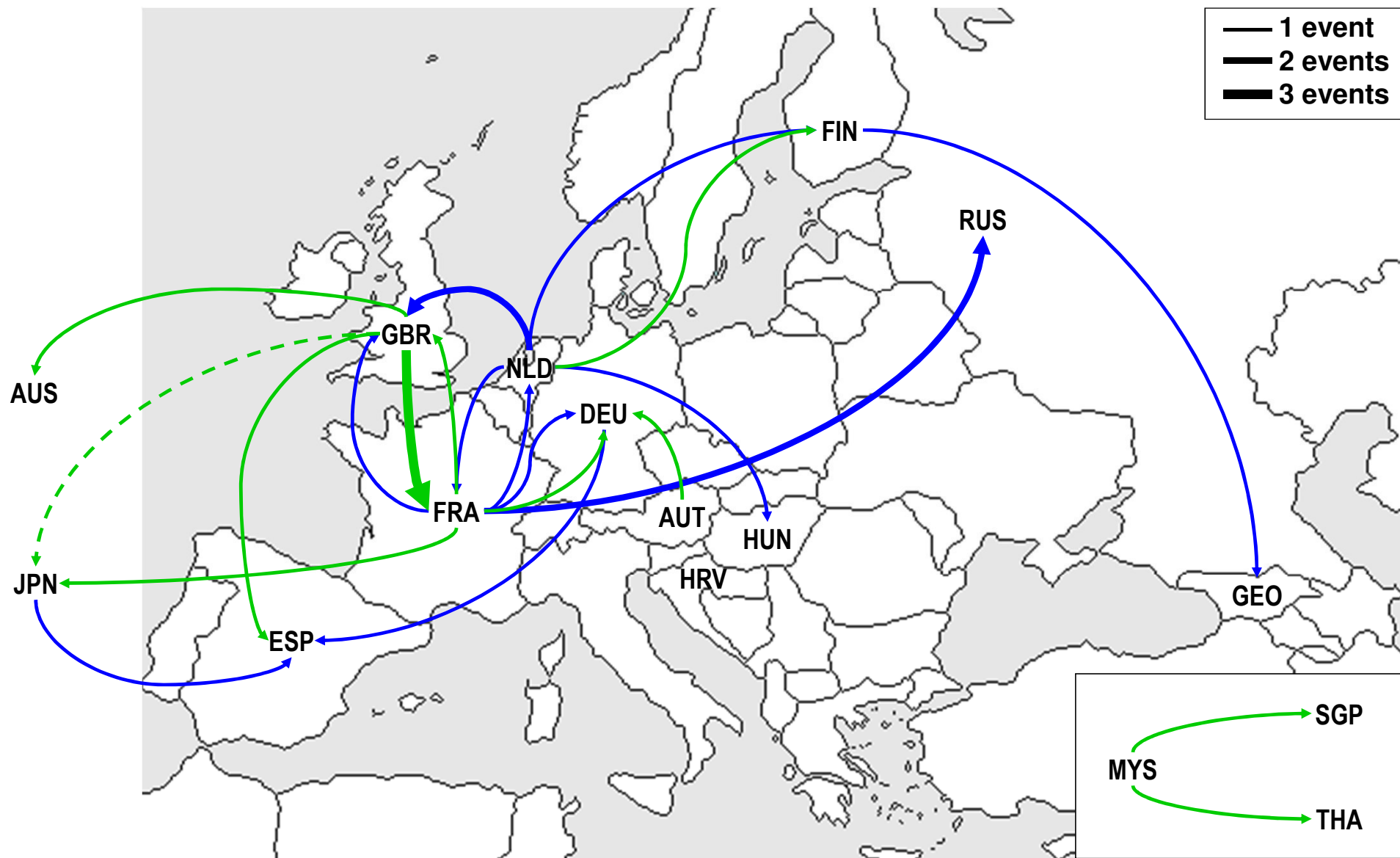
Figure 3B



**Figure 3** Spatial and temporal distribution of EV-71 lineages showing the geographic dissemination of subgenogroups C1 (A) and C2 (B)

The MCC tree was obtained from a Bayesian MCMC analysis (discrete phylogeographic model) and shows for more clarity the most relevant parts of tree topologies (time is indicated as calendar years). The FigTree program was used to display a number of information. Branches and circles at the tree nodes are colored according to the geographic location that had the highest probability. The size of each circle represents the location probability. Line width indicates the posterior probability of the corresponding lineage (thick lines indicate high posterior probability values). The tree topologies were used to determine the phylogenetic patterns (numbered sequentially) showing the most probable virus transportation events. The main features of the most probable virus transportation events (i.e. geographic location, location probability, and transportation direction) are indicated in the figure. Abbreviations used: AUS, Australia; AUT, Austria; AZE, Azerbaijan; CHE, Switzerland; DEU, Germany; ESP, Spain; FIN, Finland; FRA, France; GBR, Great Britain; GEO, Georgia; HUN, Hungary; ISL, Island; JPN, Japan; KGZ, Kyrgyzstan; LVA, Latvia; MYS, Malaysia; NLD, Netherlands; NOR, Norway; PHL, Philippines; RUS, Russia; SGP, Singapore; THA, Thailand; TWN, Taiwan; UKR, Ukraine; USA, United states of America.

Figure 4

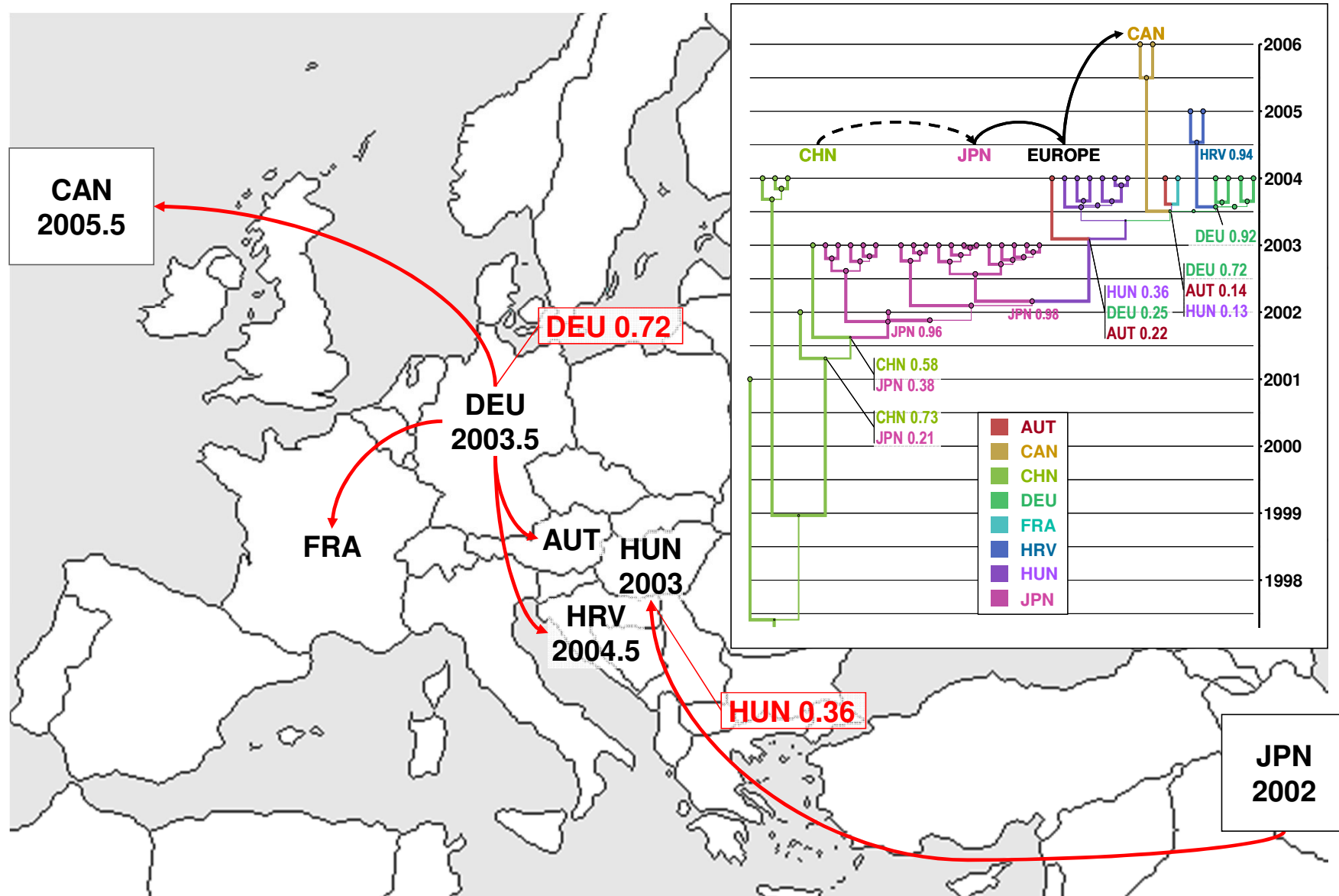


**Figure 4** Geographic dispersal of EV-A71 subgenogroups C1 (green) and C2 (blue) in Europe and virus transportation between Europe other regions

Schematic representation of the most probable virus transportation events indicated in figures 3 and 4 showing the diffusion of EV-A71 strains. Line width represents the number of transportation events inferred between two countries as indicated in the upper right scale. Abbreviations: see the legend to Figure 3.



Figure 5



**Figure 5** Bayesian analyses of the spatial and temporal distribution of an EV-A71 subgenogroup C4 lineage

The topology was drawn as described in the legend to figure 3 with the sequence sample corresponding to the subtree indicated in figure 1. The geographic locations and the location probabilities estimated with a discrete phylogeographic model are indicated for the most relevant nodes. The most probable virus transportation events are shown in the map insert. Abbreviations: see the legend to Figure 3.

# **Manuscrit n°2**

**« Enterovirus A71 Subgenotype B5,  
France, 2013 »**

**publié dans le journal « Emerging Infectious  
Diseases » sous forme d'une lettre.**

**Accepté le 9 janvier 2015.**

**Auteurs : Audrey Mirand, Lucie Molet, Chervin  
Hassel, Hélène Peigue-Lafeuille, Flore  
Rozenberg, Jean-Luc Bailly, Cécile Henquell.**

## Enterovirus A71 Subgenotype B5, France, 2013

Audrey Mirand, Lucie Molet, Chervin Hassel, Hélène Peigue-Lafeuille, Flore Rozenberg, Jean-Luc Bailly, Cécile Henquell

Author affiliations: Clermont Université, Clermont-Ferrand, France (A. Mirand, C. Hassel, H. Peigue-Lafeuille, J.-L. Bailly, C. Henquell); Centre National de Référence des Entérovirus-Parechovirus, Clermont-Ferrand (A. Mirand, H. Peigue-Lafeuille, J.-L. Bailly, C. Henquell); Hôpital Cochin, Paris, France (L. Molet, F. Rozenberg); Université Paris Descartes, Paris (L. Molet, F. Rozenberg)

DOI: <http://dx.doi.org/10.3201/eid2104.141093>

**To the Editor:** We report the detection of human enterovirus A71 (EV-A71) subgenotype B5 in France, 6 years after it was first detected in Europe. EV-A71 belongs to the *Enterovirus A* species (genus *Enterovirus*, family *Picornaviridae*) and is a major cause of hand, foot and mouth disease (HFMD), sometimes associated with severe neurologic complications (1). EV-A71 strains are classified in 6 genotypes, A–F, (2) but most of the circulating strains belong to genotypes B and C and to 11 subgenotypes (B0–B5, C1–C5) (1). Genotypes B and C have been reported in HFMD epidemics in the Asia-Pacific region; different subgenotypes cause nationwide epidemics that usually occur every 2–3 years (1). During 1963–1986 in Europe, Australia, and the United States, enterovirus infections were caused by viruses from subgenotypes B0, B1, and B2, but since 2000, infections with C1 and C2 viruses have begun to predominate (3,4). The other subgenotypes have been reported rarely in Europe; the C4b and C4a strains were identified in France, Germany, Austria, and Denmark in 2004 and 2012, respectively (4–6), and the B5 subgenotype was reported in Denmark in 2007 (7).

In November 2013, a 3-week-old boy was admitted to the emergency unit of a hospital in Compiègne, France, with a 48-hour history of fever and irritability. He was born at term after an uneventful pregnancy and delivery. On admission, he had normal vital signs. He had had contact with a cousin with oral ulcerations, but no information was available about the source of the cousin's infection.

Laboratory testing revealed moderate cytolytic hepatitis. Complete blood count results were within reference values. Cerebrospinal fluid showed pleiocytosis (38 leukocytes, 81% polymorphonuclear cells) with protein and glucose levels with reference ranges. Bacterial cultures of blood and cerebrospinal fluid were negative. Enterovirus genome was detected in serum samples and cerebrospinal fluid by reverse transcription PCR. The infant made a steady recovery and was discharged 10 days after admission,

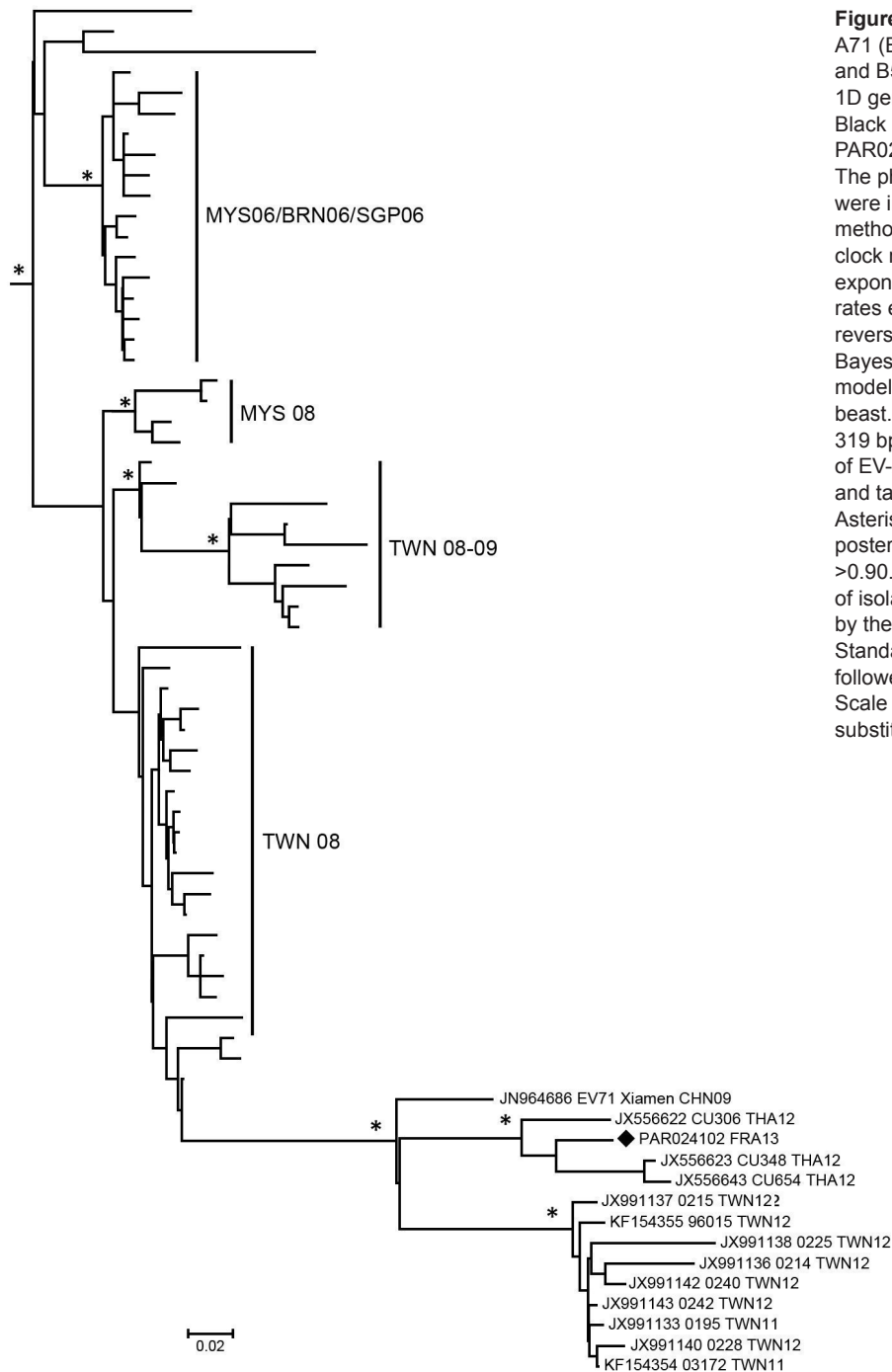
with no apparent adverse outcome. Final diagnosis was neonatal enterovirus infection with meningitis.

Genotyping was performed on the serum specimen by using seminested reverse transcription PCR amplification and sequencing of the viral protein 1 gene. Phylogenetic investigation with sequences of reference strains representing all subgenotypes indicated that the isolate from the patient, designated PAR024102\_FRA13, belonged to the EV-A71 B5 subgenotype. We investigated the putative origin of the strain by comparing 248 nonredundant complete 1D sequences of EV-A71 B5 strains following a Bayesian phylogenetic approach. PAR024103\_FRA13 shared a most recent common ancestor (posterior probability = 1) with virus strains sampled in China in 2009 and in Taiwan during the 2011–2012 outbreak, (8) but was only distantly related to them (data not shown).

Further analyses with 274 partial 1D gene sequences from GenBank (on March 19, 2014) indicated close genetic relationships (posterior probability = 1) with strains isolated in Thailand in 2012 (9) (Figure 1). The complete genome of PAR024103\_FRA13 was determined by nucleotide sequencing of 4 overlapping segments obtained by gene amplification (GenBank accession no. LK985324). Sequence comparisons were performed with 13 available EV-A71 B5 complete genomes; the virus strain isolated in France exhibited 92%–99.5% nt similarity (98.9%–99.4% aa similarity) throughout the genome.

The EV-A71 B5 subgenotype was first detected in 1999 in Malaysia and spread to several other countries in Asia during the 2000s. Outbreaks causing severe illness and deaths were reported in Japan (2003), Brunei (2006), and Taiwan (2008 and 2012) (1,8). The first detection of subgenotype B5 in Europe was associated with a recrudescence of EV-A71 infections associated with meningitis and HFMD in Denmark in 2007 (7). The overall phylogenetic data are consistent with an introduction of EV-A71 B5 in France by importation of a strain from Asia, possibly from Thailand. Transmission of EV-A71 strains has been shown to occur in Europe as discrete and temporally defined virus introductions, occasionally followed by sustained dissemination (C. Hassel, unpub. data).

The emergence of the Asiatic lineage EV-A71 C4a in Denmark in 2012 is a recent event (6). The reemergence of EV-A71 subgenotype B5 in 2008 in Taiwan resulted in the largest outbreak of EV-A71 infection in the past 11 years (8). To our knowledge, the B5 subgenotype has not previously been detected in Europe. Global herd immunity produced by circulation of the C2 genotype may protect the European population from the spread of other subgenotypes (3). However, given that most countries in Europe do not perform specific surveillance for HFMD and most enterovirus infections are asymptomatic, this particular subgenotype could be circulating more widely without detection.



**Figure.** Phylogeny of enterovirus A71 (EV-A71) subgenogroups B4 and B5 inferred with 274 partial 1D gene sequences, France. Black diamond indicates strain PAR024103\_FRA13 from this study. The phylogenetic relationships were inferred following a Bayesian method by using a relaxed molecular clock model with an uncorrelated exponential distribution of evolution rates estimated with a general time reversible substitution model and a Bayesian skyline plot as a population model (BEAST version 1.7.5; <http://beast.bio.ed.ac.uk/>). Sequences were 319 bp. For clarity, only a subtree of EV-A71 B5 sequences is shown, and taxon names are not included. Asterisks indicate nodes with posterior probability density values >0.90. Geographic origins and time of isolation of strains are indicated by the International Organization of Standardization code abbreviation followed by the year of isolation. Scale bar indicates nucleotide substitutions per site.

Enterovirus infections in neonates and infants are a frequent cause of hospitalization, which may contribute to EV-A71 detection (5). However, the development of a national syndromic surveillance targeting HFMD would enable early detection of HFMD outbreaks and any new EV-A71 subgenotype. Attention should also be paid to the potential risks of epidemic spread of EV-A71 outside Asia posed by international travelers.

**Acknowledgments**

We are grateful to Nathalie Rodde and Gwendoline Jugie for excellent technical assistance in enterovirus genotyping and complete genome sequencing. We thank Jeffrey Watts for revising the English manuscript.

The Centre National de Référence des Enterovirus-Parechovirus is supported by an annual grant from the French national public health network (Institut de Veille Sanitaire).

## References

1. Solomon T, Lewthwaite P, Perera D, Cardoso MJ, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect Dis*. 2010;10:778–90. [http://dx.doi.org/10.1016/S1473-3099\(10\)70194-8](http://dx.doi.org/10.1016/S1473-3099(10)70194-8)
2. Bessaud M, Razadindratsimandresy R, Nougairède A, Joffret ML, Deshpande JM, Dubot-Pères A, et al. Molecular comparison and evolutionary analyses of VP1 nucleotide sequences of new African human enterovirus 71 isolates reveal a wide genetic diversity. *PLoS One*. 2014;9:e90624. <http://dx.doi.org/10.1371/journal.pone.0090624>
3. van der Sanden S, van der Avoort, Lemey P, Uslu G, Koopmans M. Evolutionary trajectory of the VP1 gene of human enterovirus 71 genogroup B and C viruses. *J Gen Virol*. 2010;91:1949–58. <http://dx.doi.org/10.1099/vir.0.019695-0>
4. Mirand A, Schuffenecker I, Henquell C, Billaud G, Jugie G, Falcon D, et al. Phylogenetic evidence of a recent spread of two populations of human enterovirus 71 in European countries. *J Gen Virol*. 2010;91:2263–77. <http://dx.doi.org/10.1099/vir.0.021741-0>
5. Schuffenecker I, Henquell C, Mirand A, Coste-Burel M, Marque-Juillet S, Desbois D, et al. New introductions of enterovirus 71 subgenogroup C4, France, 2012. *Emerg Infect Dis*. 2014;20:1343–6.
6. Fischer TK, Nielsen AY, Sydenham TV, Andersen PH, Andersen B, Midgley SE. Emergence of enterovirus 71 C4a in Denmark, 2009 to 2013. *Euro Surveill*. 2014;19:20911.
7. Badran SA, Midgley S, Andersen P, Böttlinger B. Clinical and virological features of enterovirus 71 infections in Denmark, 2005 to 2008. *Scand J Infect Dis*. 2011;43:642–8. <http://dx.doi.org/10.3109/00365548.2011.577094>
8. Wu WH, Kuo TC, Lin YT, Huang SW, Liu HF, Wang J, et al. Molecular epidemiology of enterovirus 71 infection in the central region of Taiwan from 2002 to 2012. *PLoS One*. 2013;8:e83711. <http://dx.doi.org/10.1371/journal.pone.0083711>
9. Linsuwanon P, Puenpa J, Huang SW, Wang YF, Mauleekoonphairoj J, Wang JR, et al. Epidemiology and seroepidemiology of human enterovirus 71 among Thai populations. *J Biomed Sci*. 2014;21:16. <http://dx.doi.org/10.1186/1423-0127-21-16>

Address for correspondence: Audrey Mirand, CHU Clermont-Ferrand, Laboratoire de Virologie, Centre National de Référence des Entérovirus-Parechovirus–Laboratoire Associé, Clermont-Ferrand, France; email: [amirand@chu-clermontferrand.fr](mailto:amirand@chu-clermontferrand.fr)

## Avian Influenza A(H7N9) Virus Antibodies in Close Contacts of Infected Persons, China, 2013–2014

Mai-Juan Ma,<sup>1</sup> Guang-Yuan Ma,<sup>1</sup> Xiao-Xian Yang,<sup>1</sup> Shan-Hui Chen,<sup>1</sup> Gregory C. Gray, Teng Zhao, Jing Bao, Jing-Jing Zhou, Yan-Hua Qian, Bin Lu, Xia Ling,<sup>2</sup> Wu-Chun Cao<sup>2</sup>

Author affiliations: Beijing Institute of Microbiology and Epidemiology, Beijing, China (M.-J. Ma, W.-C. Cao); Wuxi Center for Disease Control and Prevention, Wuxi, China (G.-Y. Ma, S.-H. Chen, J. Bao, Y.-H. Qian, B. Lu, X. Ling); State Key

Laboratory of Pathogen and Biosecurity, Beijing (X.-X. Yang, T. Zhao, J.-J. Zhou); School of Public Health, Central South University, Changsha, China (X.-X. Yang); Duke University Medical Center, Durham, North Carolina, USA (G.C. Gray)

DOI: <http://dx.doi.org/10.3201/eid2104.141442>

**To the Editor:** From early 2013 (1) through November 2014, >460 human cases of laboratory-confirmed avian influenza A(H7N9) virus infection occurred in China. Although human-to-human transmission of subtype H7N9 virus is not common, evidence has been reported of probable transmission among several family clusters (2), between 2 household contacts (3), and between a doctor and an infected patient (4). Taken together, these observations suggest that family members, health care providers, and other close contacts (hereafter called contacts) of H7N9-infected persons may be at risk for infection.

In China, national guidelines regarding H7N9-infected patients call for observation of contacts for 7 days after exposure for signs and symptoms of infection and, if any occur, collection of throat swab specimens for testing by molecular assays (5). The guidelines do not call for serologic testing. Because human avian influenza infections may be mild or asymptomatic, we sought to determine whether serologic testing would show evidence of H7N9 virus infection among contacts of infected persons during the 2013–2014 epidemic in China. Contacts were defined in accordance with China's guidelines for prevention and control of human H7N9 virus infection (5,6). The institutional review board of Wuxi Center for Disease Control and Prevention, Wuxi, Jiangsu Province, China, reviewed and approved this study.

During the epidemic, we recruited contacts of patients in Wuxi and collected throat swab specimens when signs or symptoms of infection developed; serum samples were collected 2–3 weeks later. Swab specimens were tested for H7N9 virus by using real-time reverse transcription PCR (7). Serum samples were tested for antibodies against hemagglutinin antigens of 3 avian influenza viruses (A/Anhui/1/2013 [H7N9], A/Anhui/1/2005 [H5N1]-RG5, and A/chicken/Jiangsu/1/00 [H9N2]) (8) by using a horse erythrocyte hemagglutination inhibition (HI) assay and against the hemagglutinin antigens of 2 seasonal influenza viruses (A/California/07/2009 [H1N1] and A/Victoria/210/2009 [H3N2]) by using a turkey erythrocyte HI assay. Serum samples with HI titers  $\geq 1:40$  against H7N9 virus were confirmed positive by microneutralization assay.

Ten laboratory-confirmed human infections with H7N9 virus occurred in Wuxi during March 29, 2013–May 15, 2014. In total, 225 contacts of 7 H7N9-infected patients

<sup>1</sup>These authors contributed equally to this article.

<sup>2</sup>These senior authors contributed equally to this article.